GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2004, 22:31:40 ; Search time 346 Seconds
(without alignments)
4037.027 Million cell updates/sec

Title: US-09-522-753-5
Perfect score: 13215
Sequence: 1 MSGSTQLVAQTWRATEPRYP.....WDEEPKPLLCSQYETLSDSE 2517

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09522753/runat_15042004_143738_17456/app_query.fasta_1.2695
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=100
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09522753_@CGN_1_1_186_@runat_15042004_143738_17456 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	12922	97.8	9053	4	US-09-976-594-306	Sequence 306, App
c	2	726.5	5.5	33529	3	US-09-144-085-3	Sequence 3, Appli
	3	718	5.4	1922	2	US-08-372-652-9	Sequence 9, Appli
	4	718	5.4	1922	5	PCT-US95-16311-9	Sequence 9, Appli
c	5	689	5.2	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	6	685	5.2	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	7	670	5.1	44377	2	US-08-804-227C-7	Sequence 7, Appli
c	8	670	5.1	44377	2	US-08-804-198-1	Sequence 1, Appli
	9	668.5	5.1	4403765	3	US-09-103-840A-2	Sequence 2, Appli
	10	665	5.0	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	11	661.5	5.0	43280	2	US-08-804-227C-1	Sequence 1, Appli
c	12	652	4.9	13987	2	US-08-804-227C-13	Sequence 13, Appl
c	13	648.5	4.9	50937	3	US-09-428-517-1	Sequence 1, Appli
	14	641	4.9	28958	1	US-08-258-261B-6	Sequence 6, Appli
	15	641	4.9	28958	1	US-08-456-837-6	Sequence 6, Appli
	16	641	4.9	28958	1	US-08-457-342-6	Sequence 6, Appli
	17	641	4.9	28958	1	US-08-457-646A-6	Sequence 6, Appli
	18	641	4.9	28958	1	US-08-458-076A-6	Sequence 6, Appli
	19	641	4.9	28958	1	US-08-764-233A-4	Sequence 4, Appli
	20	641	4.9	28958	1	US-08-457-335A-6	Sequence 6, Appli
	21	641	4.9	28958	1	US-08-729-214-6	Sequence 6, Appli
	22	641	4.9	28958	3	US-09-028-934-6	Sequence 6, Appli
	23	641	4.9	49377	1	US-08-764-233A-1	Sequence 1, Appli
c	24	631.5	4.8	80161	3	US-09-036-987A-1	Sequence 1, Appli
c	25	631.5	4.8	80161	3	US-09-370-700-1	Sequence 1, Appli
c	26	631.5	4.8	80161	4	US-09-603-207-1	Sequence 1, Appli
c	27	624.5	4.7	68750	3	US-09-335-409-1	Sequence 1, Appli
c	28	624.5	4.7	68750	4	US-09-568-102-1	Sequence 1, Appli
c	29	624.5	4.7	68750	4	US-09-567-969-1	Sequence 1, Appli
c	30	624.5	4.7	68750	4	US-09-568-480-1	Sequence 1, Appli
c	31	624.5	4.7	68750	4	US-09-568-486-1	Sequence 1, Appli
c	32	624.5	4.7	68750	4	US-09-568-472-1	Sequence 1, Appli
c	33	624.5	4.7	68750	4	US-09-567-899-1	Sequence 1, Appli
c	34	622.5	4.7	71989	4	US-09-443-501A-2	Sequence 2, Appli
c	35	619	4.7	47981	4	US-09-679-279-1	Sequence 1, Appli
c	36	616.5	4.7	8438	1	US-07-945-283-1	Sequence 1, Appli
	37	614.5	4.7	15872	3	US-09-105-537-1	Sequence 1, Appli
	38	614.5	4.7	15872	4	US-09-091-609-1	Sequence 1, Appli
	39	614.5	4.7	15872	4	US-09-091-609-3	Sequence 3, Appli
c	40	606	4.6	34094	4	US-09-292-034-1	Sequence 1, Appli
	41	600	4.5	36778	3	US-09-105-537-5	Sequence 5, Appli
	42	600	4.5	38506	3	US-09-320-878-19	Sequence 19, Appl
	43	600	4.5	38506	4	US-09-141-908-1	Sequence 1, Appli
	44	600	4.5	38506	4	US-09-657-440-19	Sequence 19, Appl
c	45	596	4.5	20235	1	US-07-642-734C-3	Sequence 3, Appli
c	46	596	4.5	20235	3	US-08-439-009A-3	Sequence 3, Appli
	47	565.5	4.3	77536	4	US-09-410-551B-1	Sequence 1, Appli
	48	559	4.2	43280	2	US-08-804-227C-1	Sequence 1, Appli
	49	557.5	4.2	53500	4	US-09-266-965-76	Sequence 76, Appl
	50	550.5	4.2	6085	3	US-09-029-603-4	Sequence 4, Appli
c	51	549.5	4.2	30001	1	US-08-125-468-1	Sequence 1, Appli
c	52	549.5	4.2	30001	2	US-08-474-933-1	Sequence 1, Appli
	53	546.5	4.1	12001	1	US-08-458-568A-11	Sequence 11, Appl
c	54	544.5	4.1	77536	4	US-09-410-551B-1	Sequence 1, Appli
	55	538.5	4.1	9354	4	US-09-579-181-10	Sequence 10, Appl
	56	538	4.1	44377	2	US-08-804-227C-7	Sequence 7, Appli
	57	538	4.1	44377	2	US-08-804-198-1	Sequence 1, Appli
c	58	530	4.0	23673	4	US-09-773-816-1	Sequence 1, Appli
c	59	524.5	4.0	15872	3	US-09-105-537-1	Sequence 1, Appli
c	60	524.5	4.0	15872	4	US-09-091-609-1	Sequence 1, Appli
c	61	524.5	4.0	15872	4	US-09-091-609-3	Sequence 3, Appli
c	62	524	4.0	12001	1	US-08-458-568A-11	Sequence 11, Appl
	63	524	4.0	34094	4	US-09-292-034-1	Sequence 1, Appli
	64	523.5	4.0	13842	3	US-09-105-537-30	Sequence 30, Appl
	65	522.5	4.0	30001	1	US-08-125-468-1	Sequence 1, Appli
	66	522.5	4.0	30001	2	US-08-474-933-1	Sequence 1, Appli
c	67	520.5	3.9	13613	3	US-09-105-537-3	Sequence 3, Appli
c	68	519	3.9	11219	1	US-07-642-734C-1	Sequence 1, Appli
c	69	519	3.9	11219	3	US-08-439-009A-1	Sequence 1, Appli
	70	517.5	3.9	11220	3	US-09-105-537-32	Sequence 32, Appl
	71	516	3.9	8916	4	US-09-579-181-11	Sequence 11, Appl
c	72	510.5	3.9	32155	4	US-08-311-731A-1	Sequence 1, Appli
	73	509	3.9	18331	4	US-09-266-965-96	Sequence 96, Appl
	74	507	3.8	12249	4	US-09-266-965-74	Sequence 74, Appl
	75	506	3.8	50937	3	US-09-428-517-1	Sequence 1, Appli
	76	504.5	3.8	18034	4	US-09-266-965-75	Sequence 75, Appl
c	77	501.5	3.8	53500	4	US-09-266-965-76	Sequence 76, Appl
	78	499.5	3.8	4767	4	US-09-231-899-76	Sequence 76, Appl
c	79	497	3.8	36519	3	US-08-923-137-2	Sequence 2, Appli
c	80	493.5	3.7	7812	3	US-09-368-590-1	Sequence 1, Appli
	81	484	3.7	11495	3	US-09-056-105-9	Sequence 9, Appli
	82	481	3.6	5121	4	US-09-252-991A-15189	Sequence 15189, A
c	83	481	3.6	53526	3	US-08-658-136-2	Sequence 2, Appli
c	84	481	3.6	53577	3	US-08-658-136-1	Sequence 1, Appli
c	85	479.5	3.6	11958	3	US-09-134-246-8	Sequence 8, Appli

	86	474.5	3.6	4725	4	US-09-410-551B-24	Sequence 24, Appl
	87	474	3.6	5392	2	US-08-403-852D-1	Sequence 1, Appli
	88	474	3.6	5392	3	US-08-510-646B-1	Sequence 1, Appli
	89	474	3.6	5392	3	US-09-231-818-1	Sequence 1, Appli
	90	474	3.6	5392	4	US-09-635-359B-1	Sequence 1, Appli
	91	472.5	3.6	8438	1	US-07-945-283-1	Sequence 1, Appli
c	92	472	3.6	36778	3	US-09-105-537-5	Sequence 5, Appli
c	93	470.5	3.6	4881	4	US-09-863-859-23	Sequence 23, Appl
	94	468.5	3.5	6327	4	US-09-252-991A-14931	Sequence 14931, A
	95	468.5	3.5	11604	4	US-09-385-040-13	Sequence 13, Appl
c	96	468.5	3.5	38506	3	US-09-320-878-19	Sequence 19, Appl
c	97	468.5	3.5	38506	4	US-09-141-908-1	Sequence 1, Appli
c	98	468.5	3.5	38506	4	US-09-657-440-19	Sequence 19, Appl
	99	466.5	3.5	11604	3	US-09-385-028-13	Sequence 13, Appl
	100	466.5	3.5	11604	4	US-09-726-614-13	Sequence 13, Appl

ALIGNMENTS

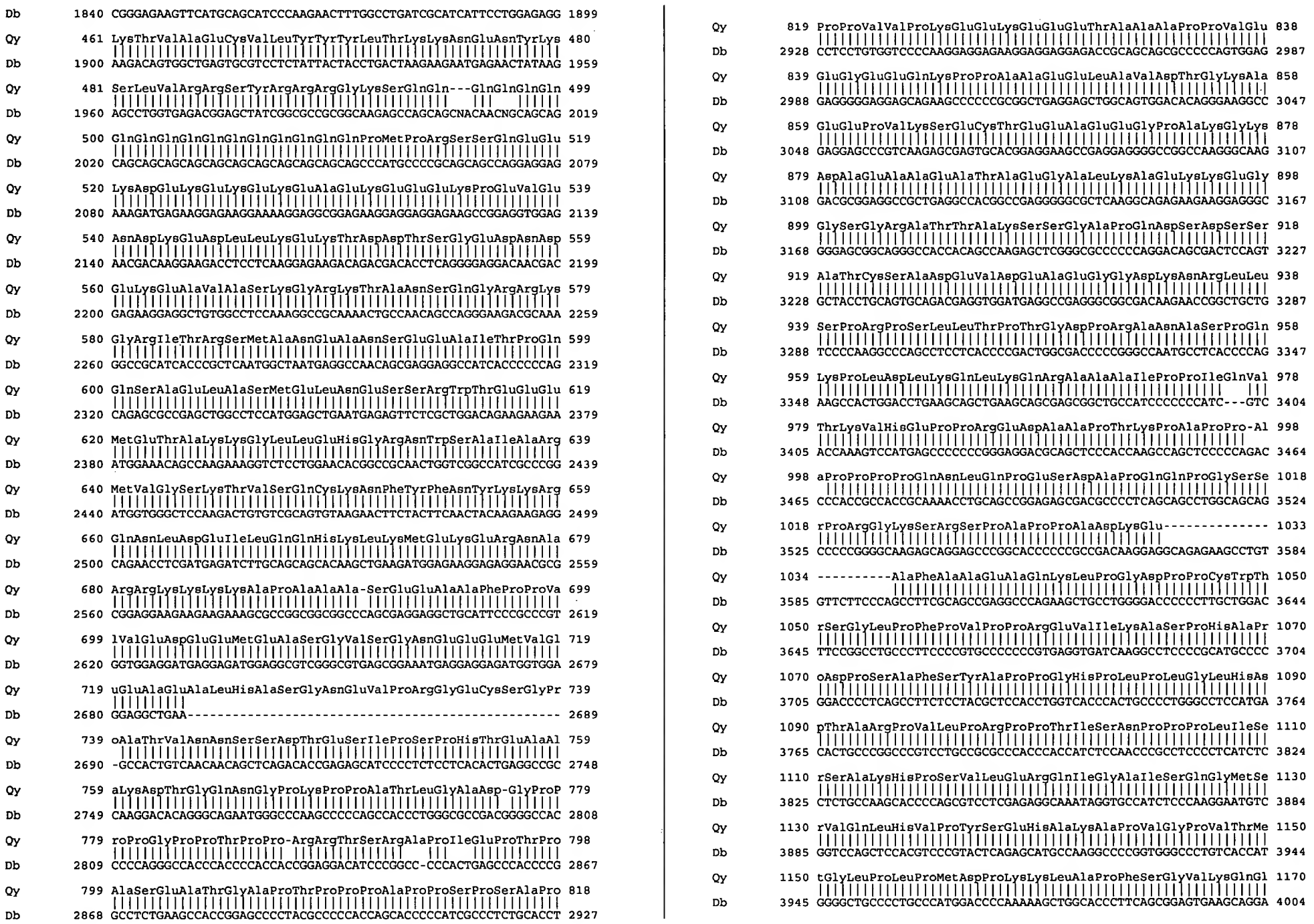
RESULT 1
US-09-976-594-306
; Sequence 306, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 306
; LENGTH: 9053
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 898877.6
; NAME/KEY: unsure
; LOCATION: 2006, 2012
; OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-306

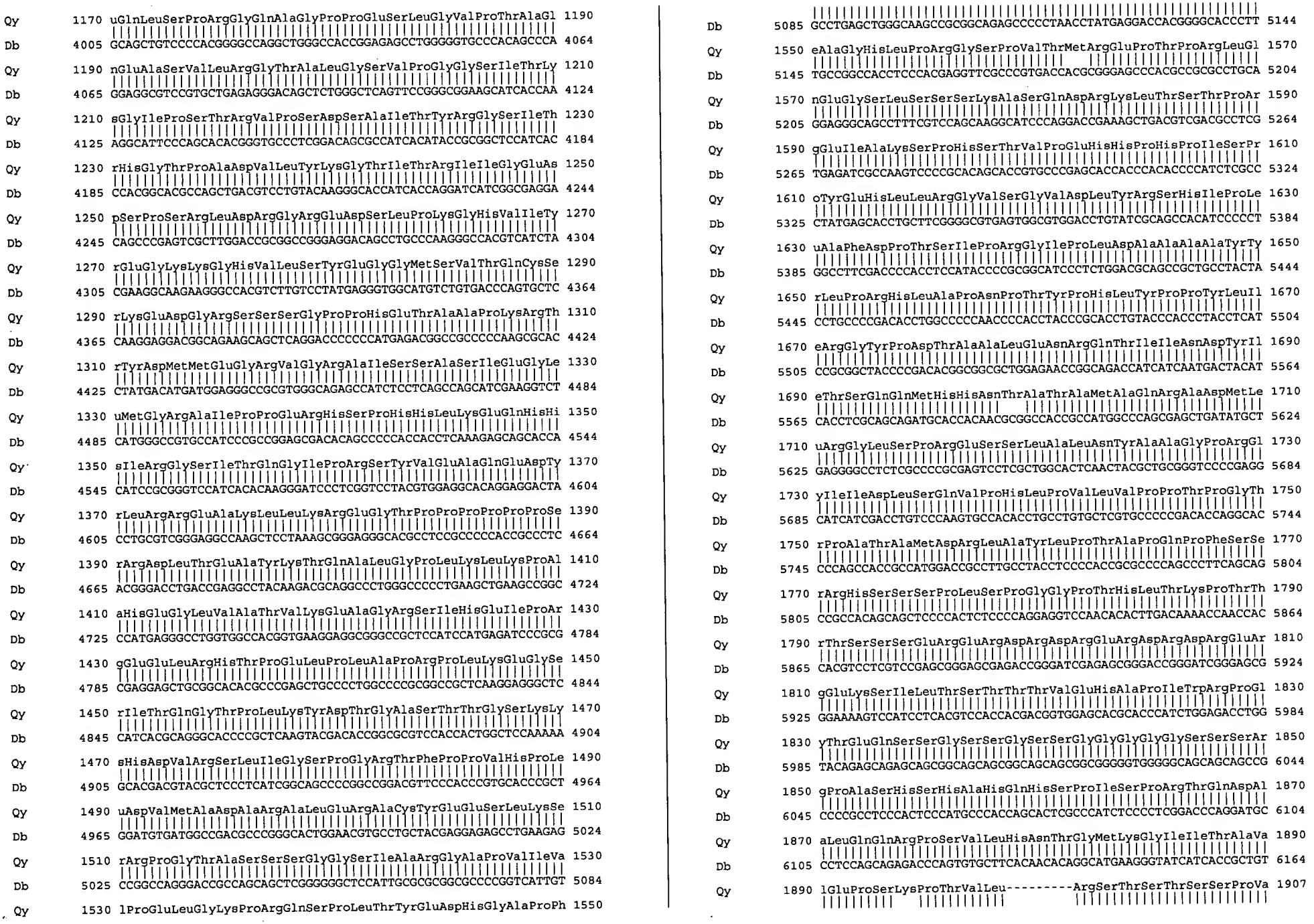
Alignment Scores:

Pred. No.:	0	Length:	9053
Score:	12922.00	Matches:	2487
Percent Similarity:	98.26%	Conservative:	2
Best Local Similarity:	98.18%	Mismatches:	10
Query Match:	97.78%	Indels:	35
DB:	4	Gaps:	5

US-09-522-753-5 (1-2517) x US-09-976-594-306 (1-9053)

```
Qy          1 MetSerGlySerThrGlnLeuValAlaGlnThrTrpArgAlaThrGluProArgTyrPro 20
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        520 ATGTCGGGATCCACACAGCCTGTGGCACAGACGTGGAGGGCCACTGAGCCCCGCTACCCG 579

Qy         21 ProHisSerLeuSerTyrProValGlnIleAlaArgThrHisThrAspValGlyLeuLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 CCCCACAGCCTTTCCTACCCAGTGCAGATCGCCCGGACGCACACGGACGTCGGGCTCCTG 639

Qy         41 GluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerProGlySerIleIleGln 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 GAGTACCAGCACCACTCCCGCGACTATGCCTCCCACCTGTCGCCCGGCTCCATCATCCAG 699

Qy         61 ProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGlyAsnGluArgSerGln 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 CCCCAGCGGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCAG 759

Qy         81 GluLeuHisLeuArgProGluSerHisSerTyrLeuProGluLeuGlyLysSerGluMet 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 GAGCTCCACCTGCGGCCAGAGTCCCACTCATACCTGCCCGAGCTGGGGAAGTCAGAGATG 819

Qy        101 GluPheIleGluSerLysArgProArgLeuGluLeuLeuProAspProLeuLeuArgPro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        820 GAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCTGCCTGACCCCCTGCTGCGACCG 879

Qy        121 SerProLeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSer 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        880 TCACCCCTGCTGGCCACGGGCCAGCCTGCGGGATCTGAAGACCTCACCAAGGACCGTAGC 939

Qy        141 LeuThrGlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeu 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        940 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCCCAGCCCCCCGCACACTGACCCTGAGCTG 999

Qy        161 GluLeuValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1000 GAGCTGGTGCCGCCACGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGACCGCGTGGAC 1059

Qy        181 ArgGluIleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1060 CGAGAGATCACCATGGTAGAGCAGCAGATCTCTAAGCTGAAGAAGAAGCAGCAACAGCTG 1119

Qy        201 GluGluGluAlaAlaLysProProGluProGluLysProValSerProProProIleGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1120 GAGGAGGAGGCTGCCAAGCCGCCCGAGCCTGAGAAGCCCGTGTCACCGCCGCCCATCGAG 1179

Qy        221 SerLysHisArgSerLeuValGlnIleIleTyrAspGluAsnArgLysLysAlaGluAla 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1180 TCGAAGCACCGCAGCCTGGTGCAGATCATCTACGACGAGAACCGGAAGAAGGCTGAAGCT 1239

Qy        241 AlaHisArgIleLeuGluGlyLeuGlyProGlnValGluLeuProLeuTyrAsnGlnPro 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1240 GCACATCGGATTCTGGAAGGCCTGGGGCCCCAGGTGGAGCTGCCGCTGTACAACCAGCCC 1299

Qy        261 SerAspThrArgGlnTyrHisGluAsnIleLysIleAsnGlnAlaMetArgLysLysLeu 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1300 TCCGACACCCGGCAGTATCATGAGAACATCAAAATAAACCAGGCGATGCGGAAGAAGCTA 1359

Qy        281 IleLeuTyrPheLysArgArgAsnHisAlaArgLysGlnTrpLysGlnLysPheCysGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db       1360 ATCTTGTACTTCAAGAGGAGGAATCACGCTCGGAAACAATGGGAGCAGAAGTTCTGCCAG 1419

Qy        301 ArgTyrAspGlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnPro 320
              ||||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db       1420 CGCTATGACCAGCTCATGGAGGCCTGGGAGAAGAAGGTGGAGCGCATCGAGAACAACCCC 1479

Qy        321 ArgArgArgAlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIle 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1480 CGGCGGCGGGCCAAGGAGAGCAAGGTGCGCGAGTACTACGAGAAGCAGTTCCCTGAGATC 1539

Qy        341 ArgLysGlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1540 CGCAAGCAGCGCGAGCTGCAGGAGCGCATGCAGAGCAGGGTGGGCCAGCGGGGCAGTGGG 1599

Qy        361 LeuSerMetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeuSer 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1600 CTGTCCATGTCGGCCGCCCGCAGCGAGCACGAGGTGTCAGAGATCATCGATGGCCTCTCA 1659

Qy        381 GluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeuTyr 400
              ||||||||||||  :::||||||||||||||||||||||||||||||||||||||||||||
Db       1660 GAGCAGGAGAACCGGAAGAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATGCTGTAC 1719

Qy        401 AspAlaAspGlnGlnArgIleLysPheIleAsnMetAsnGlyLeuMetAlaAspProMet 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1720 GACGCTGACCAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTATGGCCGACCCCATG 1779

Qy        421 LysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluThrPhe 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1780 AAGGTGTACAAAGACCGCCAGGTCATGAACATGTGGAGTGAGCAGGAGAAGGAGACCTTC 1839

Qy        441 ArgGluLysPheMetGlnHisProLysAsnPheGlyLeuIleAlaSerPheLeuGluArg 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

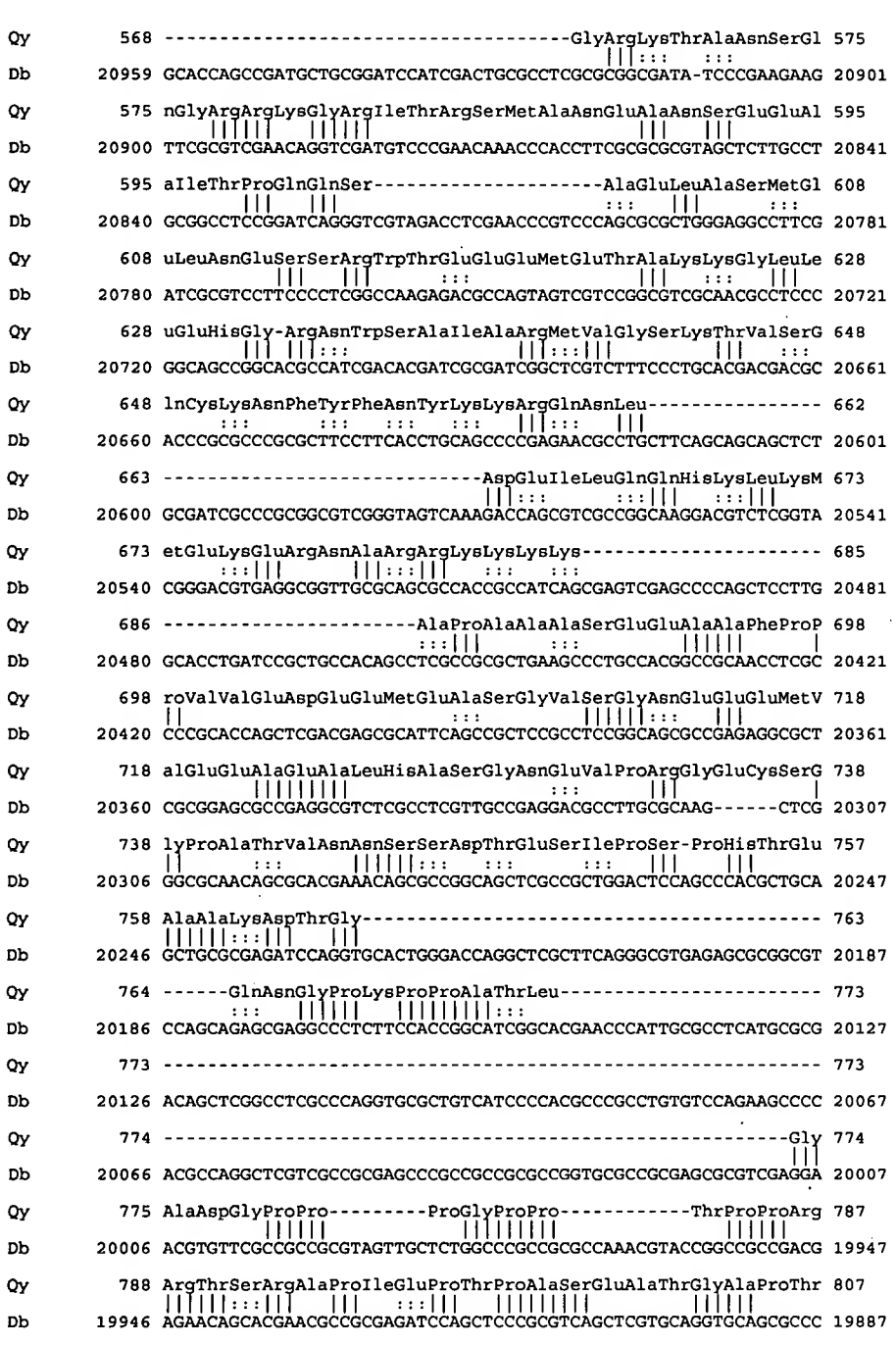

Db 1840 CGGGAGAAGTTCATGCAGCATCCCAAGAACTTTGGCCTGATCGCATCATTCCTGGAGAGG 1899
Qy 461 LysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAsnTyrLys 480
Db 1900 AAGACAGTGGCTGAGTGCGTCCTCTATTACTACCTGACTAAGAAGAATGAGAACTATAAG 1959
Qy 481 SerLeuValArgArgSerTyrArgArgArgGlyLysSerGlnGln---GlnGlnGlnGln 499
Db 1960 AGCCTGGTGAGACGGAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCNACAACNGCAGCAG 2019
Qy 500 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnProMetProArgSerSerGlnGluGlu 519
Db 2020 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCCCATGCCCCGCAGCAGCCAGGAGGAG 2079
Qy 520 LysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluValGlu 539
Db 2080 AAAGATGAGAAGGAGAAGGAAAAGGAGGCGGAGAAGGAGGAGGAGAAGCCGGAGGTGGAG 2139
Qy 540 AsnAspLysGluAspLeuLeuLysGluLysThrAspAspThrSerGlyGluAspAsnAsp 559
Db 2140 AACGACAAGGAAGACCTCCTCAAGGAGAAGACAGACGACACCTCAGGGGAGGACAACGAC 2199
Qy 560 GluLysGluAlaValAlaSerLysGlyArgLysThrAlaAsnSerGlnGlyArgArgLys 579
Db 2200 GAGAAGGAGGCTGTGGCCTCCAAAGGCCGCAAAACTGCCAACAGCCAGGGAAGACGCAAA 2259
Qy 580 GlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAlaIleThrProGln 599
Db 2260 GGCCGCATCACCCGCTCAATGGCTAATGAGGCCAACAGCGAGGAGGCCATCACCCCCCAG 2319
Qy 600 GlnSerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrpThrGluGluGlu 619
Db 2320 CAGAGCGCCGAGCTGGCCTCCATGGAGCTGAATGAGAGTTCTCGCTGGACAGAAGAAGAA 2379
Qy 620 MetGluThrAlaLysLysGlyLeuLeuGluHisGlyArgAsnTrpSerAlaIleAlaArg 639
Db 2380 ATGGAAACAGCCAAGAAAGGTCTCCTGGAACACGGCCGCAACTGGTCGGCCATCGCCCGG 2439
Qy 640 MetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLysLysArg 659
Db 2440 ATGGTGGGCTCCAAGACTGTGTCGCAGTGTAAGAACTTCTACTTCAACTACAAGAAGAGG 2499
Qy 660 GlnAsnLeuAspGluIleLeuGlnGlnHisLysLeuLysMetGluLysGluArgAsnAla 679
Db 2500 CAGAACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGGAGAAGGAGAGGAACGCG 2559
Qy 680 ArgArgLysLysLysAlaProAlaAlaAla-SerGluGluAlaAlaPheProProVa 699
Db 2560 CGGAGGAAGAAGAAGAAAGCGCCGGCGGCGGCCCAGCGAGGAGGCTGCATTCCCGCCCGT 2619
Qy 699 lValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetValGl 719
Db 2620 GGTGGAGGATGAGGAGATGGAGGCGTCGGGCGTGAGCGGAAATGAGGAGGAGATGGTGGA 2679
Qy 719 uGluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerGlyPr 739
Db 2680 GGAGGCTGAA---------- 2689
Qy 739 oAlaThrValAsnAsnSerSerAspThrGluSerIleProSerProHisThrGluAlaAl 759
Db 2690 -GCCACTGTCAACAACAGCTCAGACACCGAGAGCATCCCCTCTCCTCACACTGAGGCCGC 2748
Qy 759 aLysAspThrGlyGlnAsnGlyProLysProProAlaThrLeuGlyAlaAsp-GlyProP 779
Db 2749 CAAGGACACAGGGCAGAATGGGCCCAAGCCCCCAGCCACCCTGGGCGCCGACGGGGCCAC 2808
Qy 779 roProGlyProProThrProPro-ArgArgThrSerArgAlaProIleGluProThrPro 798
Db 2809 CCCCAGGGCCACCCACCCCACCACCGGAGGACATCCCGGCC-CCCACTGAGCCCACCCCG 2867
Qy 799 AlaSerGluAlaThrGlyAlaProThrProProProAlaProProSerProSerAlaPro 818
Db 2868 GCCTCTGAAGCCACCGGAGCCCCTACGCCCCCACCAGCACCCCCATCGCCCTCTGCACCT 2927
Qy 819 ProProValValProLysGluGluLysGluGluGluThrAlaAlaAlaProProValGlu 838
Db 2928 CCTCCTGTGGTCCCCAAGGAGGAGAAGGAGGAGGAGACCGCAGCAGCGCCCCCAGTGGAG 2987
Qy 839 GluGlyGluGluGlnLysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAla 858
Db 2988 GAGGGGGAGGAGCAGAAGCCCCCCGCGGCTGAGGAGCTGGCAGTGGACACAGGGAAGGCC 3047
Qy 859 GluGluProValLysSerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLys 878
Db 3048 GAGGAGCCCGTCAAGAGCGAGTGCACGGAGGAAGCCGAGGAGGGGCCGGCCAAGGGCAAG 3107
Qy 879 AspAlaGluAlaAlaGluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGly 898
Db 3108 GACGCGGAGGCCGCTGAGGCCACGGCCGAGGGGGCGCTCAAGGCAGAGAAGAAGGAGGGC 3167
Qy 899 GlySerGlyArgAlaThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSer 918
Db 3168 GGGAGCGGCAGGGCCACCACAGCCAAGAGCTCGGGCGCCCCCCAGGACAGCGACTCCAGT 3227
Qy 919 AlaThrCysSerAlaAspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeuLeu 938
Db 3228 GCTACCTGCAGTGCAGACGAGGTGGATGAGGCCGAGGGCGGCGACAAGAACCGGCTGCTG 3287
Qy 939 SerProArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerProGln 958
Db 3288 TCCCCAAGGCCCAGCCTCCTCACCCCGACTGGCGACCCCCGGGCCAATGCCTCACCCCAG 3347
Qy 959 LysProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGlnVal 978
Db 3348 AAGCCACTGGACCTGAAGCAGCTGAAGCAGCGAGCGGCTGCCATCCCCCCCATC---GTC 3404
Qy 979 ThrLysValHisGluProProArgGluAspAlaAlaProThrLysProAlaProPro-Al 998
Db 3405 ACCAAAGTCCATGAGCCCCCCCGGGAGGACGCAGCTCCCACCAAGCCAGCTCCCCCAGAC 3464
Qy 998 aProProProProGlnAsnLeuGlnProGluSerAspAlaProGlnGlnProGlySerSe 1018
Db 3465 CCCACCGCCACCGCAAAACCTGCAGCCGGAGAGCGACGCCCCTCAGCAGCCTGGCAGCAG 3524
Qy 1018 rProArgGlyLysSerArgSerProAlaProProAlaAspLysGlu---------- 1033
Db 3525 CCCCCGGGGCAAGAGCAGGAGCCCGGCACCCCCCGCCGACAAGGAGGCAGAGAAGCCTGT 3584
Qy 1034 ----------AlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProProCysTrpTh 1050
Db 3585 GTTCTTCCCAGCCTTCGCAGCCGAGGCCCAGAAGCTGCCTGGGGACCCCCCTTGCTGGAC 3644
Qy 1050 rSerGlyLeuProPheProValProProArgGluValIleLysAlaSerProHisAlaPr 1070
Db 3645 TTCCGGCCTGCCCTTCCCCGTGCCCCCCCGTGAGGTGATCAAGGCCTCCCCGCATGCCCC 3704
Qy 1070 oAspProSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGlyLeuHisAs 1090
Db 3705 GGACCCCTCAGCCTTCTCCTACGCTCCACCTGGTCACCCACTGCCCCTGGGCCTCCATGA 3764
Qy 1090 pThrAlaArgProValLeuProArgProProThrIleSerAsnProProProLeuIleSe 1110
Db 3765 CACTGCCCGGCCCGTCCTGCCGCGCCCACCCACCATCTCCAACCCGCCTCCCCTCATCTC 3824
Qy 1110 rSerAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIleSerGlnGlyMetSe 1130
Db 3825 CTCTGCCAAGCACCCCAGCGTCCTCGAGAGGCAAATAGGTGCCATCTCCCAAGGAATGTC 3884
Qy 1130 rValGlnLeuHisValProTyrSerGluHisAlaLysAlaProValGlyProValThrMe 1150
Db 3885 GGTCCAGCTCCACGTCCCGTACTCAGAGCATGCCAAGGCCCCGGTGGGCCCTGTCACCAT 3944
Qy 1150 tGlyLeuProLeuProMetAspProLysLysLeuAlaProPheSerGlyValLysGlnGl 1170
Db 3945 GGGGCTGCCCCTGCCCATGGACCCCAAAAAGCTGGCACCCTTCAGCGGAGTGAAGCAGGA 4004

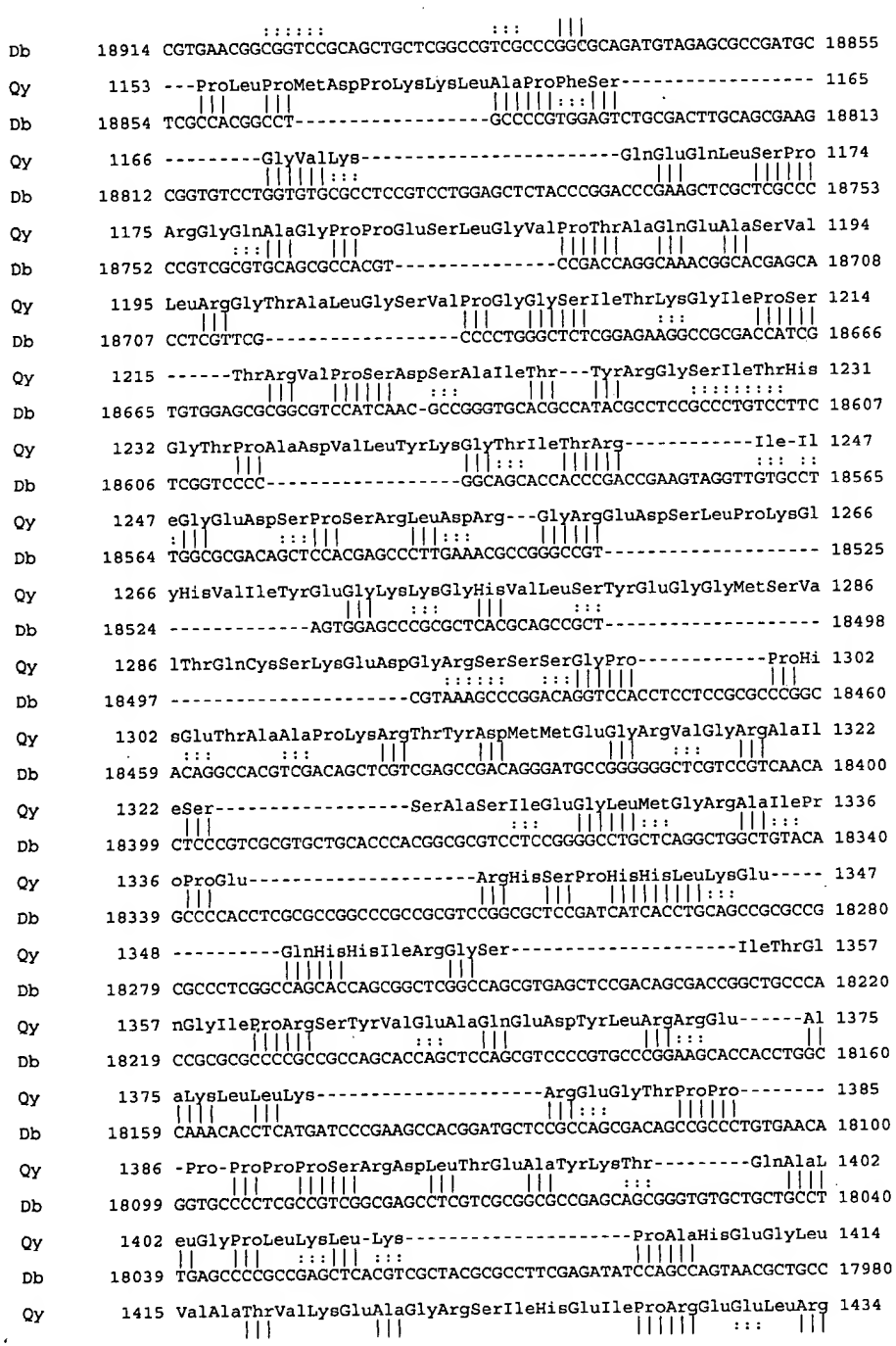

Qy 1170 uGlnLeuSerProArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGl 1190
Db 4005 GCAGCTGTCCCCACGGGGCCAGGCTGGGCCACCGGAGAGCCTGGGGGTGCCCCACAGCCCA 4064
Qy 1190 nGluAlaSerValLeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLy 1210
Db 4065 GGAGGCGTCCGTGCTGAGAGGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAA 4124
Qy 1210 sGlyIleProSerThrArgValProSerAspSerAlaIleThrTyrArgGlySerIleTh 1230
Db 4125 AGGCATTCCCAGCACACGGGTGCCCTCGGACAGCGCCATCACATACCGCGGCTCCATCAC 4184
Qy 1230 rHisGlyThrProAlaAspValLeuTyrLysGlyThrIleThrArgIleIleGlyGluAs 1250
Db 4185 CCACGGCACGCCAGCTGACGTCCTGTACAAGGGCACCATCACCAGGATCATCGGCGAGGA 4244
Qy 1250 pSerProSerArgLeuAspArgGlyArgGluAspSerLeuProLysGlyHisValIleTy 1270
Db 4245 CAGCCCGAGTCGCTTGGACCGCGGCCGGGAGGACAGCCTGCCCAAGGGCCACGTCATCTA 4304
Qy 1270 rGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerValThrGlnCysSe 1290
Db 4305 CGAAGGCAAGAAGGGCCACGTCTTGTCCTATGAGGGTGGCATGTCTGTGACCCAGTGCTC 4364
Qy 1290 rLysGluAspGlyArgSerSerSerGlyProProHisGluThrAlaAlaProLysArgTh 1310
Db 4365 CAAGGAGGACGGCAGAAGCAGCTCAGGACCCCCCCATGAGACGGCCGCCCCCAAGCGCAC 4424
Qy 1310 rTyrAspMetMetGluGlyArgValGlyArgAlaIleSerSerAlaSerIleGluGlyLe 1330
Db 4425 CTATGACATGATGGAGGGCCGCGTGGGCAGAGCCATCTCCTCAGCCAGCATCGAAGGTCT 4484
Qy 1330 uMetGlyArgAlaIleProProGluArgHisSerProHisHisLeuLysGluGlnHisHi 1350
Db 4485 CATGGGCCGTGCCATCCCGCCGGAGCGACACAGCCCCCACCACCTCAAAGAGCAGCACCA 4544
Qy 1350 sIleArgGlySerIleThrGlnGlyIleProArgSerTyrValGluAlaGlnGluAspTy 1370
Db 4545 CATCCGCGGGTCCATCACACAAGGGATCCCTCGGTCCTACGTGGAGGCACAGGAGGACTA 4604
Qy 1370 rLeuArgArgGluAlaLysLeuLeuLysArgGluGlyThrProProProProProProSe 1390
Db 4605 CCTGCGTCGGGAGGCCAAGCTCCTAAAGCGGGAGGGCACGCCTCCGCCCCCACCGCCCTC 4664
Qy 1390 rArgAspLeuThrGluAlaTyrLysThrGlnAlaLeuGlyProLeuLysLeuLysProAl 1410
Db 4665 ACGGGACCTGACCGAGGCCTACAAGACGCAGGCCCTGGGCCCCCTGAAGCTGAAGCCGGC 4724
Qy 1410 aHisGluGlyLeuValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProAr 1430
Db 4725 CCATGAGGGCCTGGTGGCCACGGTGAAGGAGGCGGGCCGCTCCATCCATGAGATCCCGCG 4784
Qy 1430 gGluGluLeuArgHisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySe 1450
Db 4785 CGAGGAGCTGCGGCACACGCCCGAGCTGCCCCTGGCCCCGCGGCCGCTCAAGGAGGGCTC 4844
Qy 1450 rIleThrGlnGlyThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLy 1470
Db 4845 CATCACGCAGGGCACCCCGCTCAAGTACGACACCGGCGCGTCCACCACTGGCTCCAAAAA 4904
Qy 1470 sHisAspValArgSerLeuIleGlySerProGlyArgThrProProProValHisProLe 1490
Db 4905 GCACGACGTACGCTCCCTCATCGGCAGCCCCGGCCGGACGTTCCCACCCGTGCACCCGCT 4964
Qy 1490 uAspValMetAlaAspAlaArgAlaLeuGluArgAlaCysTyrGluGluSerLeuLysSe 1510
Db 4965 GGATGTGATGGCCGACGCCCGGGCACTGGAACGTGCCTGCTACGAGGAGAGCCTGAAGAG 5024
Qy 1510 rArgProGlyThrAlaSerSerSerGlyGlySerIleAlaArgGlyAlaProValIleVa 1530
Db 5025 CCGGCCAGGGACCGCCAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCCGGTCATTGT 5084
Qy 1530 lProGluLeuGlyLysProArgGlnSerProLeuThrTyrGluAspHisGlyAlaProPh 1550
Db 5085 GCCTGAGCTGGGCAAGCCGCGGCAGAGCCCCCTAACCTATGAGGACCACGGGGCACCCTT 5144
Qy 1550 eAlaGlyHisLeuProArgGlySerProValThrMetArgGluProThrProArgLeuGl 1570
Db 5145 TGCCGGCCACCTCCCACGAGGTTCGCCCGTGACCACGCGGGAGCCCACGCCGCGCCTGCA 5204
Qy 1570 nGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProAr 1590
Db 5205 GGAGGGCAGCCTTTCGTCCAGCAAGGCATCCCAGGACCGAAAGCTGACGTCGACGCCTCG 5264
Qy 1590 gGluIleAlaLysSerProHisSerThrValProGluHisHisProHisProIleSerPr 1610
Db 5265 TGAGATCGCCAAGTCCCCGCACAGCACCGTGCCCGAGCACCACCCACACCCCATCTCGCC 5324
Qy 1610 oTyrGluHisLeuLeuArgGlyValSerGlyValAspLeuTyrArgSerHisIleProLe 1630
Db 5325 CTATGAGCACCTGCTTCGGGGCGTGAGTGGCGTGGACCTGTATCGCAGCCACATCCCCCT 5384
Qy 1630 uAlaPheAspProThrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTy 1650
Db 5385 GGCCTTCGACCCCACCTCCATACCCCGCGGCATCCCTCTGGACGCAGCCGCTGCCTACTA 5444
Qy 1650 rLeuProArgHisLeuAlaProAsnProThrTyrProHisLeuTyrProProTyrLeuIl 1670
Db 5445 CCTGCCCCGACACCTGGCCCCCAACCCCACCTACCCGCACCTGTACCCACCCTACCTCAT 5504
Qy 1670 eArgGlyTyrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIl 1690
Db 5505 CCGCGGCTACCCCGACACGGCGGCGCTGGAGAACCGGCAGACCATCATCAATGACTACAT 5564
Qy 1690 eThrSerGlnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLe 1710
Db 5565 CACCTCGCAGCAGATGCACCACAACGCGGCCACCGCCATGGCCCAGCGAGCTGATATGCT 5624
Qy 1710 uArgGlyLeuSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGl 1730
Db 5625 GAGGGGCCTCTCGCCCCGCGAGTCCTCGCTGGCACTCAACTACGCTGCGGGTCCCCGAGG 5684
Qy 1730 yIleIleAspLeuSerGlnValProHisLeuProValLeuValProProThrProGlyTh 1750
Db 5685 CATCATCGACCTGTCCCAAGTGCCACACCTGCCTGTGCTCGTGCCCCCGACACCAGGCAC 5744
Qy 1750 rProAlaThrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSe 1770
Db 5745 CCCAGCCACCGCCATGGACCGCCTTGCCTACCTCCCCACCGCGCCCCAGCCCTTCAGCAG 5804
Qy 1770 rArgHisSerSerSerProLeuSerProGlyGlyProThrHisLeuThrLysProThrTh 1790
Db 5805 CCGCCACAGCAGCTCCCCACTCTCCCCAGGAGGTCCAACACACTTGACAAAACCAACCAC 5864
Qy 1790 rThrSerSerSerGluArgGluArgAspArgAspArgGluArgAspArgAspArgGluAr 1810
Db 5865 CACGTCCTCGTCCGAGCGGGAGCGAGACCGGGATCGAGAGCGGGACCGGGATCGGGAGCG 5924
Qy 1810 gGluLysSerIleLeuThrSerThrThrThrValGluHisAlaProIleTrpArgProGl 1830
Db 5925 GGAAAAGTCCATCCTCACGTCCACCACGACGGTGGAGCACGCACCCATCTGGAGACCTGG 5984
Qy 1830 yThrGluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGlySerSerSerAr 1850
Db 5985 TACAGAGCAGAGCAGCGGCAGCAGCGGCAGCAGCGGCGGGGGTGGGGGCAGCAGCAGCCG 6044
Qy 1850 gProAlaSerHisSerHisAlaHisGlnHisSerProIleSerProArgThrGlnAspAl 1870
Db 6045 CCCCGCCTCCCACTCCCATGCCCACCAGCACTCGCCCATCTCCCCTCGGACCCAGGATGC 6104
Qy 1870 aLeuGlnGlnArgProSerValLeuHisAsnThrGlyMetLysGlyIleIleThrAlaVa 1890
Db 6105 CCTCCAGCAGAGACCCAGTGTGCTTCACAACACAGGCATGAAGGGTATCATCACCGCTGT 6164
Qy 1890 lGluProSerLysProThrValLeu---------ArgSerThrSerThrSerSerProVa 1907

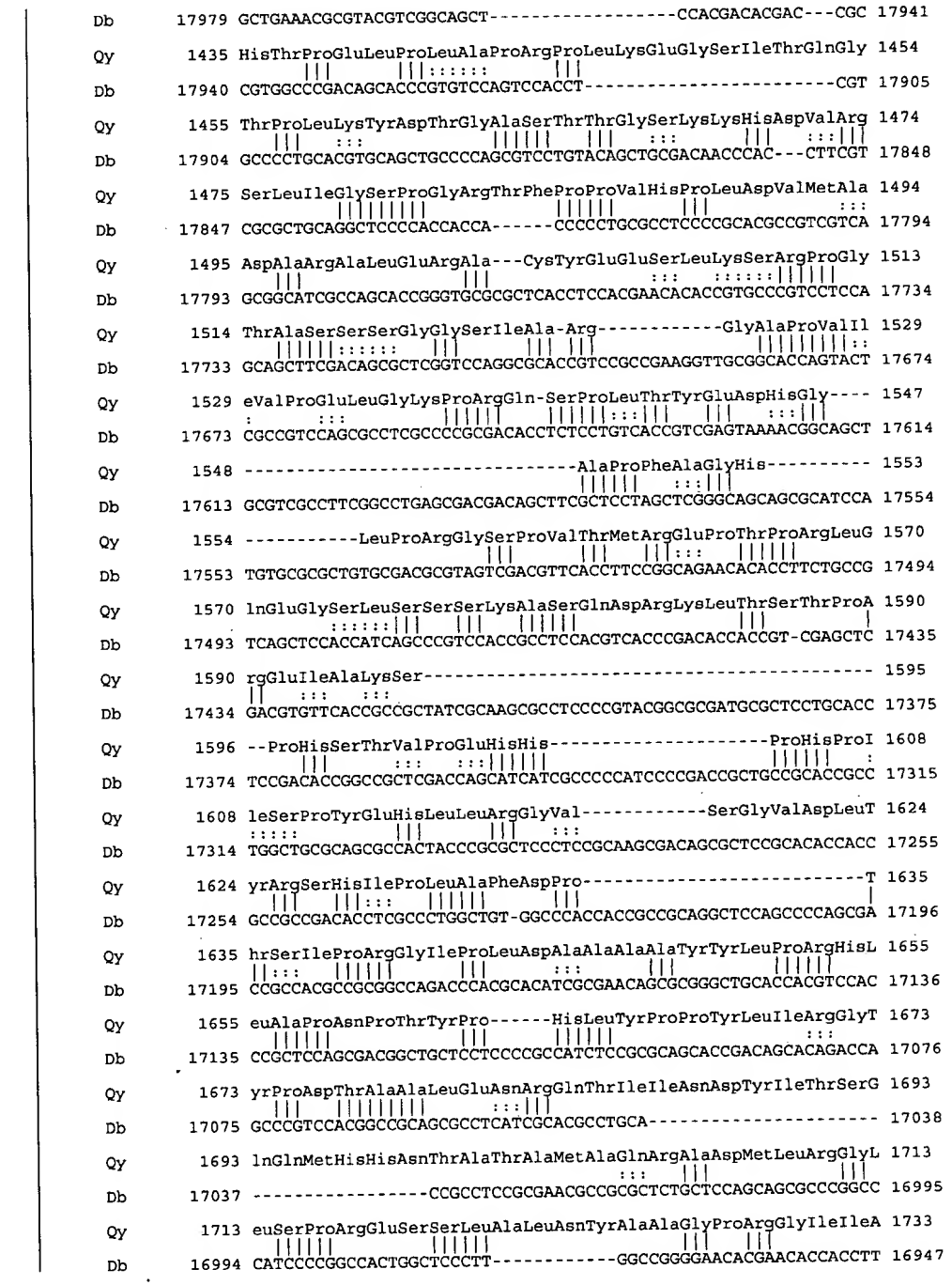
```
Db       6165 GGAGCCCAGCACGCCCACGGTCCTGAGGTGGGCCAGGTCCACCTCCACCTCCTCACCCGT 6224

Qy       1907 lArgProAlaAlaThrPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGl 1927
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6225 TCGCCCGGCTGCCACATTCCCACCTGCCACCCACTGCCCACTGGGCGGCACCCTCGATGG 6284

Qy       1927 yValTyrProThrLeuMetGluProValLeuLeuProLysGluAlaProArgValAlaAr 1947
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6285 GGTCTACCCTACCCTCATGGAGCCCGTCTTGCTGCCCAAGGAGGCCCCCCGGGTCGCCCG 6344

Qy       1947 gProGluArgProArgAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSe 1967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6345 GCCAGAGCGGCCCCGAGCAGACACCGGCCATGCCTTCCTCGCCAAGCCCCCAGCCCGCTC 6404

Qy       1967 rGlyLeuGluProAlaSerSerProSerLysGlySerGluProArgProLeuValProPr 1987
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6405 CGGGCTGGAGCCCGCCTCCTCCCCCAGCAAGGGCTCGGAGCCCCGGCCCCTAGTGCCTCC 6464

Qy       1987 oValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAl 2007
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6465 TGTCTCTGGCCACGCCACCATCGCCCGCACCCCTGCGAAGAACCTCGCACCTCACCACGC 6524

Qy       2007 aSerProAspProProAlaProProAlaSerAlaSerAspProHisArgGluLysThrGl 2027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6525 CAGCCCGGACCCGCCGGCGCCACCTGCCTCGGCCTCGGACCCGCACCGGGAAAAGACTCA 6584

Qy       2027 nSerLysProPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSe 2047
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6585 AAGTAAACCCTTTTCCATCCAGGAACTGGAACTCCGTTCTCTGGGTTACCACGGCAGCAG 6644

Qy       2047 rTyrSerProGluGlyValGluProValSerProValSerSerProSerLeuThrHisAs 2067
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6645 CTACAGCCCCGAAGGGGTGGAGCCCGTCAGCCCTGTGAGCTCACCCAGTCTGACCCACGA 6704

Qy       2067 pLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuAr 2087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6705 CAAGGGGCTCCCCAAGCACCTGGAAGAGCTCGACAAGAGCCACCTGGAGGGGGAGCTGCG 6764

Qy       2087 gProLysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuAr 2107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6765 GCCCAAGCAGCCAGGCCCCGTGAAGCTTGGCGGGGAGGCCGCCCACCTCCCACACCTGCG 6824

Qy       2107 gProLeuProGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLy 2127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6825 GCCGCTGCCTGAGAGCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGCCCCAGGGGTCAA 6884

Qy       2127 sGlyHisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTy 2147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6885 AGGTCACCAGCGGGTGGTCACCCTGGCCCAGCACATCAGTGAGGTCATCACACAGGACTA 6944

Qy       2147 rThrArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPhePr 2167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6945 CACCCGGCACCACCCACAGCAGCTCAGCGCACCCCTGCCCGCCCCCCCTCTACTCCTTCCC 7004

Qy       2167 oGlyAlaSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProPr 2187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7005 TGGGGCCAGCTGCCCCGTCCTGGACCTCCGCCGCCCACCCAGTGACCTCTACCTCCCGCC 7064

Qy       2187 oProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArgSerPr 2207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7065 CCCGGACCATGGTGCCCCGGCCCGTGGCTCCCCCCACAGCGAAGGGGGCAAGAGGTCTCC 7124

Qy       2207 oGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGlyIleGluProValSerPr 2227
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7125 AGAGCCAAACAAGACGTCGGTCTTGGGTGGTGGTGAGGACGGTATTGAACCTGTGTCCCC 7184

Qy       2227 oProGluGlyMetThrGluProGlyHisSerArgSerAlaValTyrProLeuLeuTyrAr 2247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7185 ACCGGAGGGCATGACGGAGCCAGGGCACTCCCGGAGTGCTGTGTACCCGCTGCTGTACCG 7244

Qy       2247 gAspGlyGluGlnThrGluProSerArgMetGlySerLysSerProGlyAsnThrSerGl 2267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7245 GGATGGGGAACAGACGGAGCCCAGCAGGATGGGCTCCAAGTCTCCAGGCAACACCAGCCA 7304

Qy       2267 nProProAlaPhePheSerLysLeuThrGluSerAsnSerAlaMetValLysSerLysLy 2287
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7305 GCCGCCAGCCTTCTTCAGCAAGCTGACCGAGAGCAACTCCGCCATGGTCAAGTCCAAGAA 7364

Qy       2287 sGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSe 2307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7365 GCAAGAGATCAACAAGAAGCTGAACACCCACAACCGGAATGAGCCTGAATACAATATCAG 7424

Qy       2307 rGlnProGlyThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTy 2327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7425 CCAGCCTGGGACGGAGATCTTCAATATGCCCGCCATCACCGGAACAGGCCTTATGACCTA 7484

Qy       2327 rArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleAr 2347
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7485 TAGAAGCCAGGCGGTGCAGGAACATGCCAGCACCAACATGGGGCTGGAGGCCATAATTAG 7544

Qy       2347 gLysAlaLeuMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAl 2367
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7545 AAAGGCACTCATGGGTAAATATGACCAGTGGGAAGAGTCCCCGCCGCTCAGCGCCAATGC 7604

Qy       2367 aPheAsnProLeuAsnAlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAs 2387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7605 TTTTAACCCTCTGAATGCCAGTGCCAGCCTGCCCGCTGCTATGCCCATAACCGCTGCTGA 7664

Qy       2387 pGlyArgSerAspHisThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGl 2407
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7665 CGGACGGAGTGACCACACACTCACCTCGCCAGGTGGCGGCGGGAAGGCCAAGGTCTCTGG 7724

Qy       2407 yArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspArgPr 2427
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7725 CAGACCCAGCAGCCGAAAAGCCAAGTCCCCGGCCCCGGGCCTGGCATCTGGGGACCGGCC 7784

Qy       2427 oProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAs 2447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7785 ACCCTCTGTCTCCTCAGTGCACTCGGAGGGAGACTGCAACCGCCGGACGCCGCTCACCAA 7844

Qy       2447 nArgValTrpGluAspArgProSerSerAlaGlySerThrProPheProTyrAsnProLe 2467
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7845 CCGCGTGTGGGAGGACAGGCCCTCGTCCGCAGGTTCCACGCCATTCCCCTACAACCCCCT 7904

Qy       2467 uIleMetArgLeuGlnAlaGlyValMetAlaSerProProProProGlyLeuProAlaGl 2487
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7905 GATCATGCGGCTGCAGGCGGGTGTCATGGCTTCCCCACCCCCACCGGGCCTCCCCGCGGG 7964

Qy       2487 ySerGlyProLeuAlaGlyProHisHisAlaTrpAspGluGluProLysProLeuLeuCy 2507
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       7965 CAGCGGGCCCCTCGCTGGCCCCCACCACGCCTGGGACGAGGAGCCCAAGCCACTGCTCTG 8024

Qy       2507 sSerGlnTyrGluThrLeuSerAspSerGlu 2517
              |||||||||||||||||||||||||||||||
Db       8025 CTCGCAGTACGAGACACTCTCCGACAGCGAG 8055

RESULT 2
US-09-144-085-3/c
; Sequence 3, Application US/09144085
; Patent No. 6280999
; GENERAL INFORMATION:
;  APPLICANT: Gustafsson, Claes
;  APPLICANT: Betlach, Mary C.
;  APPLICANT: Ashley, Gary
;  APPLICANT: Julien, Bryan
;  APPLICANT: Ziermann, Rainer
;  TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
;  TITLE OF INVENTION: THEREFOR
;  FILE REFERENCE: 30062-20020.20
;  CURRENT APPLICATION NUMBER: US/09/144,085
;  CURRENT FILING DATE: 1998-08-31
;  EARLIER APPLICATION NUMBER: 09/010,809
;  EARLIER FILING DATE: 1998-01-22
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
```

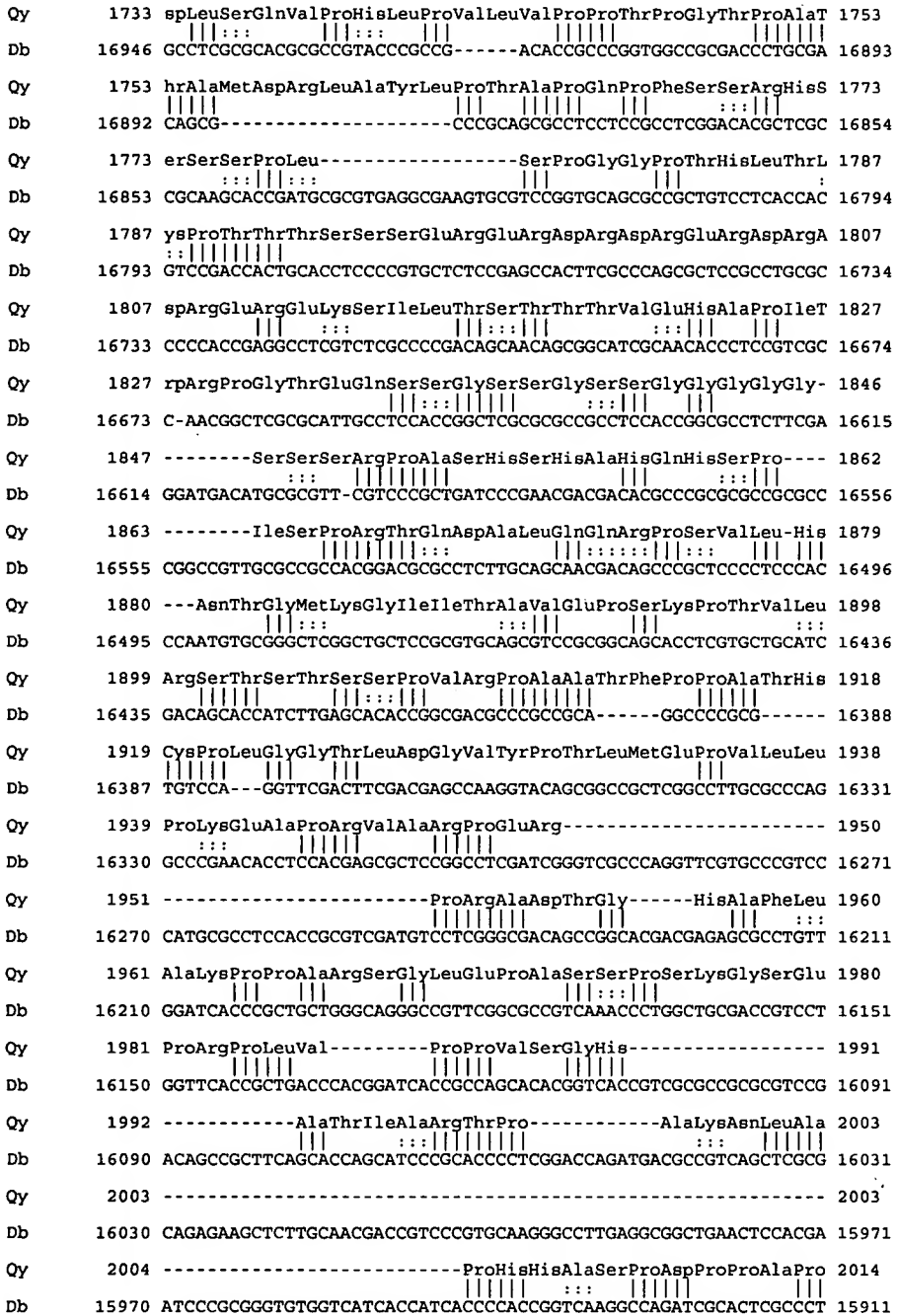
```
;   LENGTH: 33529
;   TYPE: DNA
;   ORGANISM: Sorangium cellulosum
US-09-144-085-3

Alignment Scores:
Pred. No.:                7.28e-21      Length:          33529
Score:                    726.50        Matches:         674
Percent Similarity:       32.36%        Conservative:    314
Best Local Similarity:    22.08%        Mismatches:      1199
Query Match:              5.50%         Indels:          873
DB:                       3             Gaps:            137

US-09-522-753-5 (1-2517) x US-09-144-085-3 (1-33529)

Qy        17 ProArgTyrProPro---HisSerLeuSerTyrProValGlnIleAlaArgThrHisThr 35
             ||||||  ||||||     ||||||     |||     :::|||
Db     22686 CCGAGGCCTCCACCGACACCATCGCTCCCCCGGCCTCGCACCCTTGCA------TCAGCC 22633

Qy        36 AspValGlyLeuLeuGluTyrGlnHisHisSerArgAspTyrAlaSerHisLeuSerPro 55
             |||:::|||                  |||     ||||||
Db     22632 GACCTCGGG-----------------------CGCACACTAGCTTCGCTGCGTCCGCAA 22597

Qy        56 GlySerIleIleGlnProGlnArgArgArgProSerLeuLeuSerGluPheGlnProGly 75
             ||||||         |||     :::|||  :::|||||||||         |||
Db     22596 GGCTCAGCACACCCGCCACATGCGCAGCGCTCAGCTCTCCTATCGAGT---GCCCCAGCA 22540

Qy        76 AsnGluArgSerGln-----------------------GluLeuHis----------- 83
             :::|||:::|||                         :::   |||
Db     22539 GCGCAGCGGGCTCAGCCCCCACGACTCCCACTGACGGTACAGCGCCACTTCCAGCGCAAA 22480

Qy        84 -----------------LeuArgPro---GluSerHisSerTyrLeuProGluLeuGly 96
                              |||||||||   :::     |||:::|||        |||
Db     22479 CAGCCCGGGCTGCGTGTACTCCGTCCGCTCCAGCAACGCTCCTTCCTCGCTGCCCGCGGC 22420

Qy        97 LysSer---------GluMetGluPheIleGluSerLysArgProArgLeuGluLeuLeu 113
             :::                       :::|||   :::     ||||||     |||
Db     22419 CGCGAACACCACCTCTCTCAACCCACGGTCGAGATACGCGTCCAGCGCCTCGCACACCTC 22360

Qy       114 ProAsp--------------------------------ProLeuLeuArgProSerPro 122
             :::                                   ||||||  |||  |||
Db     22359 GTCGAACGCCGCACGGAACACGGGGTACACTTCGTAAAGCCTCTTCCCCATCCCGAGCCG 22300

Qy       123 LeuLeuAlaThrGlyGlnProAlaGlySerGluAspLeuThrLysAspArgSerLeuThr 142
             |||  |||                                        |||
Db     22299 CTGGCTGCCCTG-------------------------------CCCCGTGAACAGCAC 22273

Qy       143 GlyLysLeuGluProValSerProProSerProProHisThrAspProGluLeuGluLeu 162
             ||||||  |||  |||          |||  |||                    :::
Db     22272 CGCAAGCTT---CCCGCCTCGCTTCGCCGTCCCGCTCAC---------CACCGCCGCATC 22225

Qy       163 ValProProArgLeuSerLysGluGluLeuIleGlnAsnMetAspArgValAspArgGlu 182
             |||||||||      :::  :::  :::  :::                 |||:::
Db     22224 CGGCCGCCCCGACGACAGCGCGCGAAGACCCTCCACAGCTCCCGCAGCGCTCGCCGCAAG 22165

Qy       183 IleThrMetValGluGlnGlnIleSerLysLeuLysLysLysGlnGlnGlnLeuGluGlu 202
                        ::::::     ::::::::::::::::
Db     22164 CACCGACGCCCGCGACTCGAAGTGCGTCCGGTGCAGCGCCGCCGT-----------GCG 22117

Qy       203 GluAlaAlaLysProProGluProGlu---------LysProValSerProProProIle 219
                ::::||||||                  :::|||:::   |||||||||
Db     22116 CACCACGTCCGACCACCCCACCTCCCCGTGCTCTTCCAGCCACTTCGCCCACCGCCCCGC 22057

Qy       220 GluSerLysHisArgSerLeuValGlnIleIleTyrAspGluAsn------ArgLysLys 237
                ||||||  ||||||                 :::       |||::::::
Db     22056 CTGCGCATTCACCGCGGCCTCGTCTCGACCCGACAGCACCAGCGGCAATAGCGCAGGCGC 21997

Qy       238 AlaGluAlaAlaHisArgIleLeuGluGlyLeuGly------------------------ 249
                      |||:::      ||||||
Db     21996 CGCCTCGGCCTCGACAGGCTCGCGCCGCGCCTCGGCCGGCGCTTCTTCGAGGATGATATG 21937

Qy       249 ------------------------------------------------------------ 249
```

```
Db     21936 CGCGTTCGTCCCGCTGATCCCGAACGACGACACGCCGGCACGACGGACCCGGCCGTTGCG 21877

Qy       250 ---ProGlnValGluLeuProLeuTyrAsnGlnProSerAspThrArgGlnTyrHisGlu 268
                |||:::      |||           ||||||:::           :::
Db     21876 CCGCCACGGACGCGCCTCTTGCAGCAGCGACAGCCCGCT---------TCCTTCCCACCC 21826

Qy       269 AsnIleLysIleAsn-----------GlnAlaMetArgLysLysLeuIleLeuTyrPhe 284
                |||::::::                |||  :::|||:::  |||:::|||:::
Db     21825 AATGTGCGGGCTCGGCTGCTCCGCGTGCAGCGTCTTCGGCAGCACCTCGTGCTGCATCGA 21766

Qy       285 LysArgArg---AsnHisAlaArgLysGlnTrpLysGlnLysPheCysGlnArgTyrAsp 303
                :::       :::|||||||||
Db     21765 CAGCACCATCTTGATCACGCCCGC---------------------------------- 21742

Qy       304 GlnLeuMetGluAlaLeuGluLysLysValGluArgIleGluAsnAsnProArgArgArg 323
                                                        :::  |||||||||
Db     21741 ----------------------------------------CACGCCCGCCGCCGC 21727

Qy       324 AlaLysGluSerLysValArgGluTyrTyrGluLysGlnPheProGluIleArgLys--- 342
                :::  |||:::||||||                    :::|||:::
Db     21726 CTGCGCATGTCCGAGGTTCGA-------------------------CTTCGACGATCC 21694

Qy       343 ------GlnArgGluLeuGlnGluArgMetGlnSerArgValGlyGlnArgGlySerGly 360
                    ||||||  |||         :::::|||:::     ::::::  ||||||
Db     21693 CAGGTACAGCGGACGCTCGGGGCTACGCTCCGGTCCAAACACCTCCGCCAGCGCTCCGGC 21634

Qy       361 LeuSer---MetSerAlaAlaArgSerGluHisGluValSerGluIleIleAspGlyLeu 379
                |||    :::|||  ||||||:::                    :::|||  |||
Db     21633 CTCGATCGGGTCTCCAAGGCTCGTGCCCGTACCGTGCGCCTCCACCGCGTCGATGTCCTC 21574

Qy       380 SerGluGlnGluAsnLeuGluLysGlnMetArgGlnLeuAlaValIleProProMetLeu 399
                   |||                  ||||||                |||:::|||
Db     21573 GGGCGACAGACCACA-----------CGACGACAGCGCCTGCCGGATCACCCGCTGCTG 21526

Qy       400 TyrAspAlaAspGlnGlnArg-IleLysPheIleAsnMetAsnGlyLeuMetAlaAspPr 419
                |||  ::::::|||       ::::::     :::         |||:::||
Db     21525 GGCAGGGCCGTTCGGCGCCGTCAGACCCTGGCTGCGACCGTCCTGGTTCACCGCAGAGCC 21466

Qy       419 oMetLysValTyrLysAspArgGlnValMetAsnMetTrpSerGluGlnGluLysGluTh 439
                |          |||                 |||  |||      :::
Db     21465 ACGGATCACCCCCAGCACACGGTCGCCGTCGCGCTGCGCGTCAGA-CAGCCGCTTCAA-- 21409

Qy       439 rPheArgGluLysPheMetGlnHisProLysAsnPheGly---LeuIleAlaSerPheLe 458
                         |||||||||        |||      |||      ||
Db     21408 ---------------CAACAGCATCCCGCAACCCTCGGCCCAGCCCGCGCCGTCCGCCT 21365

Qy       458 uGluArgLysThrValAlaGluCysValLeuTyrTyrTyrLeuThrLysLysAsnGluAs 478
                |  |||:::  :::|||      |||      :::  |||           |||
Db     21364 GCACCGAGAAGCTCTTGCACCGACCGTCGCGGGACATCCCCTT-----------GAGAC 21317

Qy       478 nTyrLysSerLeuValArgArg------SerTyrArgArgArgGlyLysSerGlnGlnGl 496
                :::          :::|||      :::::  |||           |||:::||
Db     21316 GGCTGAACTCCACGAACAGCGCGGGGGTGCTCATCACCGTCACCCCGCCGGCCAGCGCCA 21257

Qy       496 n-----------------GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln-- 509
                |                |||::::::  :::|||  |||:::  |||::::::
Db     21256 GGTCGCATTCGCCCTGGCGCAGCGCCGTGCACGCCAGGTGCAGCGACACCAGCGACGACG 21197

Qy       510 --------------------------------GlnProMetProArgSerSerGlnGl 518
                                             ||||||     |||
Db     21196 AGCACGCCGTGTCCACCGTCATCGCTGGGCCCTGCAGCCCCAG---CGCGTAGGCCACCC 21140

Qy       518 uGluLysAspGluLysGluLysGluLysGluAlaGluLysGluGluGluLysProGluVa 538
                :::::::::      :::      :::|||           :::      :::::
Db     21139 GGCCCGAGAGGATGCTCGCCGCGCTCCCGATGCCCCGGTACCCGTCCAGCTCCTTCAGAT 21080

Qy       538 lGluAsnAspLysGluAspLeuLeuLys---GluLysThrAsp---AspThrSerGlyGl 556
                :::::::   |||      :::::::::      :::|||         :::   |||
Db     21079 CGACAGTATGAAGAGCACCGTAGTCCGAGCCCATCGAACCCAGGTACACTCCGGTGGCGC 21020

Qy       556 uAspAsnAspGluLysGluAlaValAlaSerLys------------------------- 567
                         :::   |||   |||
Db     21019 TCCCGCTCAGCGCCGACGGGCGCACGCCGGCACGCTCCAATGCCTCCCACGCCGTCTCCA 20960
```


Qy 568 ----------GlyArgLysThrAlaAsnSerGl 575
Db 20959 GCACCAGCCGATGCTGCGGATCCATCGACTGCGCCTCGCGCGGCGATA-TCCCGAAGAAG 20901
Qy 575 nGlyArgArgLysGlyArgIleThrArgSerMetAlaAsnGluAlaAsnSerGluGluAl 595
Db 20900 TTCGCGTCGAACAGGTCGATGTCCCGAACAAACCCACCTTCGCGCGCGTAGCTCTTGCCT 20841
Qy 595 aIleThrProGlnGlnSer----------AlaGluLeuAlaSerMetGl 608
Db 20840 GCGGCCTCCGGATCAGGGTCGTAGACCTCGAACCCGTCCCAGCGCGCTGGGAGGCCTTCG 20781
Qy 608 uLeuAsnGluSerSerArgTrpThrGluGluGluMetGluThrAlaLysLysGlyLeuLe 628
Db 20780 ATCGCGTCCTTCCCCTCGGCCAAGAGACGCCAGTAGTCGTCCGGCGTCGCAACGCCTCCC 20721
Qy 628 uGluHisGly-ArgAsnTrpSerAlaIleAlaArgMetValGlySerLysThrValSerG 648
Db 20720 GGCAGCCGGCACGCCATCGACACGATCGCGATCGGCTCGTCTTTCCCTGCACGACGACGC 20661
Qy 648 lnCysLysAsnPheTyrPheAsnTyrLysLysArgGlnAsnLeu---------- 662
Db 20660 ACCCGCGCCCGCGCTTCCTTCACCTGCAGCCCCGAGAACGCCTGCTTCAGCAGCAGCTCT 20601
Qy 663 ----------AspGluIleLeuGlnGlnHisLysLeuLysM 673
Db 20600 GCGATCGCCCGCGGCGTCGGGTAGTCAAAGACCAGCGTCGCCGGCAAGGACGTCTCGGTA 20541
Qy 673 etGluLysGluArgAsnAlaArgArgLysLysLysLys---------- 685
Db 20540 CGGGACGTGAGGCGGTTGCGCAGCGCCACCGCCATCAGCGAGTCGAGCCCCAGCTCCTTG 20481
Qy 686 ----------AlaProAlaAlaAlaSerGluGluAlaAlaPheProP 698
Db 20480 GCACCTGATCCGCTGCCACAGCCTCGCCGCGCTGAAGCCCTGCCACGGCCGCAACCTCGC 20421
Qy 698 roValValGluAspGluGluMetGluAlaSerGlyValSerGlyAsnGluGluGluMetV 718
Db 20420 CCCGCACCAGCTCGACGAGCGCATTCAGCCGCTCCGCCTCCGGCAGCGCCGAGAGGCGCT 20361
Qy 718 alGluGluAlaGluAlaLeuHisAlaSerGlyAsnGluValProArgGlyGluCysSerG 738
Db 20360 CGCGGAGCGCCGAGGCGTCTCGCCTCGTTGCCGAGGACGCCTTGCGCAAG------CTCG 20307
Qy 738 lyProAlaThrValAsnAsnSerSerAspThrGluSerIleProSer-ProHisThrGlu 757
Db 20306 GGCGCAACAGCGCACGAAACAGCGCCGGCAGCTCGCCGCTGGACTCCAGCCCACGCTGCA 20247
Qy 758 AlaAlaLysAspThrGly---------- 763
Db 20246 GCTGCGCGAGATCCAGGTGCACTGGGACCAGGCTCGCTTCAGGGCGTGAGAGCGCGGCGT 20187
Qy 764 ------GlnAsnGlyProLysProProAlaThrLeu---------- 773
Db 20186 CCAGCAGAGCGAGGCCCTCTTCCACCGGCATCGGCACGAACCCATTGCGCCTCATGCGCG 20127
Qy 773 ---------- 773
Db 20126 ACAGCTCGGCCTCGCCCAGGTGCGCTGTCATCCCCACGCCCGCCTGTGTCCAGAAGCCCC 20067
Qy 774 ----------Gly 774
Db 20066 ACGCCAGGCTCGTCGCCGCGAGCCCGCCGCCGCGCCGGTGCGCCGCGAGCGCGTCGAGGA 20007
Qy 775 AlaAspGlyProPro---------ProGlyProPro----------ThrProProArg 787
Db 20006 ACGTGTTCGCCGCCGCGTAGTTGCTCTGGCCCGCCGCGCCAAACGTACCGGCCGCCGACG 19947
Qy 788 ArgThrSerArgAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThr 807
Db 19946 AGAACAGCACGAACGCCGCGAGATCCAGCTCCCGCGTCAGCTCGTGCAGGTGCAGCGCCC 19887
Qy 808 ProProProAlaProProSerPro---------SerAlaProProProValValProLys 824
Db 19886 CGTCCACCTTCGGCGCCAGCACCCGCGAGAGGCGCTCGGCCGTCTGGGCGGTGAGCACGC 19827
Qy 825 GluGluLysGluGluGluThrAlaAlaAlaProProValGluGluGlyGluGluGln--- 843
Db 19826 CGTCGTCGAGCACGCCGGCCAGGTGCAGCACCGCGCTCAGCGGACGCGCCGCGTCGATGC 19767
Qy 844 LysProProAlaAlaGluGluLeuAlaValAspThrGlyLysAlaGluGluProValLys 863
Db 19766 CGGCCAGCACACGCGCGACCTCCTCCCGCTTCGACACGTCGCACGCAGCCACCGTCACCG 19707
Qy 864 SerGluCysThrGluGluAlaGluGluGlyProAlaLysGlyLysAspAlaGluAlaAla 883
Db 19706 TCTCGGCGCCGAGCTCCTCGAGTGACTGCACGAGCTCGCGGGCCCCGGGCGCCTCCAGCC 19647
Qy 884 GluAlaThrAlaGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAla 903
Db 19646 CGC----------GACGCGACGTCAGCACAAGGTGCCTCACCCCGT--- 19611
Qy 904 ThrThrAlaLysSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAla 923
Db 19610 ------GCGCGCGCACCAGGTGCGCCGCGATCGCCTGACCCAGCTCCCCTGTCCCGCCGG 19557
Qy 924 AspGluValAspGluAlaGluGlyGlyAspLysAsnArgLeu----------Leu 938
Db 19556 TCACCAGCACCGTGCCCGCAGGGTCCAGCTCGCGGGCTCCTCGGGTGAGCTCTTCCGCTG 19497
Qy 939 SerPro---ArgProSerLeuLeuThrProThrGlyAspProArgAlaAsnAlaSerPro 957
Db 19496 CCGCCTGCGCGCGCACCAGGCGCGCGGCCAGCGCAGCGCCCCCGCGCAGCGCAAGCTCCG 19437
Qy 958 GlnLysProLeuAspLeuLysGlnLeuLysGlnArgAlaAlaAlaIleProProIleGln 977
Db 19436 GCTCCGCCG----------CCGTCGCCAGCGCCCGCGCCAGCAGCCCAG 19398
Qy 978 ValThrLysValHisGluProProArgGluAspAla---------- 989
Db 19397 CGTCCACGGGCTCGGTCCCCACATCGATCAAGCGCAGCCGGCGCTCGGGGTGCTCGCTGC 19338
Qy 990 AlaProThrLysProAlaProProAla------ProProProProGlnAsnLeuGlnPro 1007
Db 19337 GCGCCGCGCGAACAAGCCCCCACAGCGGCGCGCGCGCCAGATCCTGGACGGCGTCGTCCG 19278
Qy 1008 GluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSerProAla 1027
Db 19277 GCGCCGCGCCCACCGCGCCCCGCGTCACCCACACCAGCTCGACAGCGTCGAGCCGCGGCT 19218
Qy 1028 ProProAlaAspLysGluAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProPro 1047
Db 19217 CCGACAGCCACGCTTGCAGCAGCGACAGCG----------CCTGCCTCGTCGCCT 19173
Qy 1048 CysTrpThrSerGlyLeuProPheProValProProArgGluValIleLysAlaSerPro 1067
Db 19172 ---------CGTGGGACAACGACGCCACTCCCGACTGACTCTGTCCTAAAGCTGTCGTGT 19122
Qy 1068 HisAlaProAspProSerAlaPheSerTyrAlaProProGly---------- 1081
Db 19121 CGACCGCCACCCGCGCAGGCGCGCTCGCGCCCTGCTCGAGGCGCGCAACCAATGCTTCGA 19062
Qy 1082 ---------HisProLeuProLeuGlyLeuHisAspThrAlaArgProValLeuProArg 1098
Db 19061 GATCGGCAATCGCCTCCGCCCTCAGGGCTT----------CGGCCAGCCTGCCTCGTC 19014
Qy 1099 ProProThrIleSerAsnProProProLeuIleSerSerAlaLysHisProSerValLeu 1118
Db 19013 CCTCCGCTGCACCGACGACCACCA----------GCGAGCCCGCCTCCAGGGGGGCTG 18966
Qy 1119 GluArgGlnIleGlyAlaIleSerGlnGlyMetSerValGlnLeuHisValProTyrSer 1138
Db 18965 CCGCGA---------GCTCCACAGGCTGGAAGCTCACCTGATACATATGTTGGCCACCAG 18915
Qy 1139 GluHisAlaLysAlaProValGlyProValThrMetGlyLeu---------- 1152


Db 18914 CGTGAACGGCGGTCCGCAGCTGCTCGGCCGTCGCCCGGCGCAGATGTAGAGCGCCGATGC 18855
Qy 1153 ---ProLeuProMetAspProLysLysLeuAlaProPheSer---------- 1165
Db 18854 TCGCCACGGCCT----------GCCCCGTGGAGTCTGCGACTTGCAGCGAAG 18813
Qy 1166 ---------GlyValLys----------GlnGluGlnLeuSerPro 1174
Db 18812 CGGTGTCCTGGTGTGCGCCTCCGTCCTGGAGCTCTACCCGGACCCGAAGCTCGCTCGCCC 18753
Qy 1175 ArgGlyGlnAlaGlyProProGluSerLeuGlyValProThrAlaGlnGluAlaServal 1194
Db 18752 CCGTCGCGTGCAGCGCCACGT----------CCGACCAGGCAAACGGCACGAGCA 18708
Qy 1195 LeuArgGlyThrAlaLeuGlySerValProGlyGlySerIleThrLysGlyIleProSer 1214
Db 18707 CCTCGTTCG----------CCCCTGGGCTCTCGGAGAAGGCCGCGACCATCG 18666
Qy 1215 ------ThrArgValProSerAspSerAlaIleThr---TyrArgGlySerIleThrHis 1231
Db 18665 TGTGGAGCGCGGCGTCCATCAAC-GCCGGGTGCACGCCATACGCCTCCGCCCTGTCCTTC 18607
Qy 1232 GlyThrProAlaAspValLeuTyrLysGlyThrIleThrArg----------Ile-Il 1247
Db 18606 TCGGTCCCC----------GGCAGCACCACCCGACCGAAGTAGGTTGTGCCT 18565
Qy 1247 eGlyGluAspSerProSerArgLeuAspArg---GlyArgGluAspSerLeuProLysGl 1266
Db 18564 TGGCGCGACAGCTCCACGAGCCCTTGAAACGCCGGGCCGT---------- 18525
Qy 1266 yHisValIleTyrGluGlyLysLysGlyHisValLeuSerTyrGluGlyGlyMetSerVa 1286
Db 18524 ----------AGTGGAGCCCGCGCTCACGCAGCCGCT---------- 18498
Qy 1286 lThrGlnCysSerLysGluAspGlyArgSerSerSerGlyPro----------ProHi 1302
Db 18497 ----------CGTAAAGCCCGGACAGGTCCACCTCCTCCGCGCCCGGC 18460
Qy 1302 sGluThrAlaAlaProLysArgThrTyrAspMetMetGluGlyArgValGlyArgAlaIl 1322
Db 18459 ACAGGCCACGTCGACAGCTCGTCGAGCCGACAGGGATGCCGGGGGGCTCGTCCGTCAACA 18400
Qy 1322 eSer----------SerAlaSerIleGluGlyLeuMetGlyArgAlaIlePr 1336
Db 18399 CTCCCGTCGCGTGCTGCACCCACGGCGCGTCCTCCGGGGCCTGCTCAGGCTGGCTGTACA 18340
Qy 1336 oProGlu----------ArgHisSerProHisHisLeuLysGlu----- 1347
Db 18339 GCCCCACCTCGCGCCGGCCCGCCGCGTCCGGCGCTCCGATCATCACCTGCAGCCGCGCCG 18280
Qy 1348 ----------GlnHisHisIleArgGlySer----------IleThrGl 1357
Db 18279 CGCCCTCGGCCAGCACCAGCGGCTCGGCCAGCGTGAGCTCCGACAGCGACCGGCTGCCCA 18220
Qy 1357 nGlyIleProArgSerTyrValGluAlaGlnGluAspTyrLeuArgArgGlu------Al 1375
Db 18219 CCGCGCGCCCCGCCGCCAGCACCAGCTCCAGCGTCCCCGTGCCCGGAAGCACCACCTGGC 18160
Qy 1375 aLysLeuLeuLys----------ArgGluGlyThrProPro-------- 1385
Db 18159 CAAACACCTCATGATCCCGAAGCCACGGATGCTCCGCCAGCGACAGCCGCCCTGTGAACA 18100
Qy 1386 -Pro-ProProProSerArgAspLeuThrGluAlaTyrLysThr---------GlnAlaL 1402
Db 18099 GGTGCCCCTCGCCGTCGGCGAGCCTCGTCGCGGCGCCGAGCAGCGGGTGTGCTGCTGCCT 18040
Qy 1402 euGlyProLeuLysLeu-Lys----------ProAlaHisGluGlyLeu 1414
Db 18039 TGAGCCCCGCCGAGCTCACGTCGCTACGCGCCTTCGAGATATCCAGCCAGTAACGCTGCC 17980
Qy 1415 ValAlaThrValLysGluAlaGlyArgSerIleHisGluIleProArgGluGluLeuArg 1434


Db 17979 GCTGAAACGCGTACGTCGGCAGCT----------CCACGACACGAC---CGC 17941
Qy 1435 HisThrProGluLeuProLeuAlaProArgProLeuLysGluGlySerIleThrGlnGly 1454
Db 17940 CGTGGCCCGACAGCACCCGTGTCCAGTCCACCT----------CGT 17905
Qy 1455 ThrProLeuLysTyrAspThrGlyAlaSerThrThrGlySerLysLysHisAspValArg 1474
Db 17904 GCCCCTGCACGTGCAGCTGCCCCAGCGTCCTGTACAGCTGCGACAACCCAC---CTTCGT 17848
Qy 1475 SerLeuIleGlySerProGlyArgThrProProValHisProLeuAspValMetAla 1494
Db 17847 CGCGCTGCAGGCTCCCCACCACCA------CCCCCTGCGCCTCCCCGCACGCCGTCGTCA 17794
Qy 1495 AspAlaArgAlaLeuGluArgAla---CysTyrGluGluSerLeuLysSerArgProGly 1513
Db 17793 GCGGCATCGCCAGCACCGGGTGCGCGCTCACCTCCACGAACACACCGTGCCCGTCCTCCA 17734
Qy 1514 ThrAlaSerSerSerGlyGlySerIleAla-Arg----------GlyAlaProValIl 1529
Db 17733 GCAGCTTCGACAGCGCTCGGTCCAGGCGCACCGTCCGCCGAAGGTTGCGGCACCAGTACT 17674
Qy 1529 eValProGluLeuGlyLysProArgGln-SerProLeuThrTyrGluAspHisGly---- 1547
Db 17673 CGCCGTCCAGCGCCTCGCCCCGCGACACCTCTCCTGTCACCGTCGAGTAAAACGGCAGCT 17614
Qy 1548 ----------AlaProPheAlaGlyHis---------- 1553
Db 17613 GCGTCGCCTTCGGCCTGAGCGACGACAGCTTCGCTCCTAGCTCGGGCAGCAGCGCATCCA 17554
Qy 1554 ----------LeuProArgGlySerProValThrMetArgGluProThrProArgLeuG 1570
Db 17553 TGTGCGCGCTGTGCGACGCGTAGTCGACGTTCACCTTCCGGCAGAACACACCTTCTGCCG 17494
Qy 1570 lnGluGlySerLeuSerSerSerLysAlaSerGlnAspArgLysLeuThrSerThrProA 1590
Db 17493 TCAGCTCCACCATCAGCCCGTCCACCGCCTCCACGTCACCCGACACCACCGT-CGAGCTC 17435
Qy 1590 rgGluIleAlaLysSer---------- 1595
Db 17434 GACGTGTTCACCGCCGCTATCGCAAGCGCCTCCCCGTACGGCGCGATGCGCTCCTGCACC 17375
Qy 1596 --ProHisSerThrValProGluHisHis----------ProHisProI 1608
Db 17374 TCCGACACCGGCCGCTCGACCAGCATCATCGCCCCCATCCCCGACCGCTGCCGCACCGCC 17315
Qy 1608 leSerProTyrGluHisLeuLeuArgGlyVal----------SerGlyValAspLeuT 1624
Db 17314 TGGCTGCGCAGCGCCACTACCCGCGCTCCCTCCGCAAGCGACAGCGCTCCGCACACCACC 17255
Qy 1624 yrArgSerHisIleProLeuAlaPheAspPro----------T 1635
Db 17254 GCCGCCGACACCTCGCCCTGGCTGT-GGCCCACCACCGCCGCAGGCTCCAGCCCCAGCGA 17196
Qy 1635 hrSerIleProArgGlyIleProLeuAspAlaAlaAlaAlaTyrTyrLeuProArgHisL 1655
Db 17195 CCGCCACGCCGCGGCCAGACCCACGCACATCGCGAACAGCGCGGGCTGCACCACGTCCAC 17136
Qy 1655 euAlaProAsnProThrTyrPro------HisLeuTyrProProTyrLeuIleArgGlyT 1673
Db 17135 CCGCTCCAGCGACGGCTGCTCCTCCCCGCCATCTCCGCGCAGCACCGACAGCACAGACCA 17076
Qy 1673 yrProAspThrAlaAlaLeuGluAsnArgGlnThrIleIleAsnAspTyrIleThrSerG 1693
Db 17075 GCCCGTCCACGGCCGCAGCGCCTCATCGCACGCCTGCA---------- 17038
Qy 1693 lnGlnMetHisHisAsnThrAlaThrAlaMetAlaGlnArgAlaAspMetLeuArgGlyL 1713
Db 17037 ----------CCGCCTCCGCGAACGCCGCGCTCTGCTCCAGCAGCGCCCGGCC 16995
Qy 1713 euSerProArgGluSerSerLeuAlaLeuAsnTyrAlaAlaGlyProArgGlyIleIleA 1733
Db 16994 CATCCCCGGCCACTGGCTCCCTT----------GGCCGGGGAACACGAACACCACCTT 16947


Qy 1733 spLeuSerGlnValProHisLeuProValLeuValProProThrProGlyThrProAlaT 1753
Db 16946 GCCTCGCGCACGCGCCGTACCCGCCG------ACACCGCCCGGTGGCCGCGACCCTGCGA 16893
Qy 1753 hrAlaMetAspArgLeuAlaTyrLeuProThrAlaProGlnProPheSerSerArgHisS 1773
Db 16892 CAGCG--------------------CCCGCAGCGCCTCCTCCGCCTCGGACACGCTCGC 16854
Qy 1773 erSerSerProLeu------------------SerProGlyGlyProThrHisLeuThrL 1787
Db 16853 CGCAAGCACCGATGCGCGTGAGGCGAAGTGCGTCCGGTGCAGCGCCGCTGTCCTCACCAC 16794
Qy 1787 ysProThrThrThrSerSerSerGluArgGluArgAspArgAspArgGluArgAspArgA 1807
Db 16793 GTCCGACCACTGCACCTCCCCGTGCTCTCCGAGCCACTTCGCCCAGCGCTCCGCCTGCGC 16734
Qy 1807 spArgGluArgGluLysSerIleLeuThrSerThrThrThrValGluHisAlaProIleT 1827
Db 16733 CCCCACCGAGGCCTCGTCTCGCCCCGACAGCAACAGCGGCATCGCAACACCCTCCGTCGC 16674
Qy 1827 rpArgProGlyThrGluGlnSerSerGlySerSerGlySerSerGlyGlyGlyGlyGly- 1846
Db 16673 C-AACGGCTCGCGCATTGCCTCCACCGGCTCGCGCGCCGCCTCCACCGGCGCCTCTTCGA 16615
Qy 1847 --------SerSerSerArgProAlaSerHisSerHisAlaHisGlnHisSerPro---- 1862
Db 16614 GGATGACATGCGCGTT-CGTCCCGCTGATCCCGAACGACGACACGCCCGCGCGCCGCGCC 16556
Qy 1863 --------IleSerProArgThrGlnAspAlaLeuGlnGlnArgProSerValLeu-His 1879
Db 16555 CGGCCGTTGCGCCGCCACGGACGCGCCTCTTGCAGCAACGACAGCCCGCTCCCCTCCCAC 16496
Qy 1880 ---AsnThrGlyMetLysGlyIleIleThrAlaValGluProSerLysProThrValLeu 1898
Db 16495 CCAATGTGCGGGCTCGGCTGCTCCGCGTGCAGCGTCCGCGGCAGCACCTCGTGCTGCATC 16436
Qy 1899 ArgSerThrSerThrSerSerProValArgProAlaAlaThrPheProProAlaThrHis 1918
Db 16435 GACAGCACCATCTTGAGCACACCGGCGACGCCCGCCGCA------GGCCCCGCG------ 16388
Qy 1919 CysProLeuGlyGlyThrLeuAspGlyValTyrProThrLeuMetGluProValLeuLeu 1938
Db 16387 TGTCCA---GGTTCGACTTCGACGAGCCAAGGTACAGCGGCCGCTCGGCCTTGCGCCCAG 16331
Qy 1939 ProLysGluAlaProArgValAlaArgProGluArg------------------------ 1950
Db 16330 GCCCGAACACCTCCACGAGCGCTCCGGCCTCGATCGGGTCGCCCAGGTTCGTGCCCGTCC 16271
Qy 1951 ------------------------ProArgAlaAspThrGly------HisAlaPheLeu 1960
Db 16270 CATGCGCCTCCACCGCGTCGATGTCCTCGGGCGACAGCCGGCACGACGAGAGCGCCTGTT 16211
Qy 1961 AlaLysProProAlaArgSerGlyLeuGluProAlaSerSerProSerLysGlySerGlu 1980
Db 16210 GGATCACCCGCTGCTGGGCAGGGCCGTTCGGCGCCGTCAAACCCTGGCTGCGACCGTCCT 16151
Qy 1981 ProArgProLeuVal---------ProProValSerGlyHis------------------ 1991
Db 16150 GGTTCACCGCTGACCCACGGATCACCGCCAGCACACGGTCACCGTCGCGCCGCGCGTCCG 16091
Qy 1992 ------------AlaThrIleAlaArgThrPro------------AlaLysAsnLeuAla 2003
Db 16090 ACAGCCGCTTCAGCACCAGCATCCCGCACCCCTCGGACCAGATGACGCCGTCAGCTCGCG 16031
Qy 2003 -- 2003
Db 16030 CAGAGAAGCTCTTGCAACGACCGTCCCGTGCAAGGGCCTTGAGGCGGCTGAACTCCACGA 15971
Qy 2004 -------------------------ProHisHisAlaSerProAspProProAlaPro 2014
Db 15970 ATCCCGCGGGTGTGGTCATCACCATCACCCCACCGGTCAAGGCCAGATCGCACTCGCCCT 15911
Qy 2015 ProAla--------------SerAlaSerAspProHisArgGluLysThrGlnSerLys 2029
Db 15910 GACGCAGGGCCGTGCACGCCAGGTGCAGCGACACCAGCGACGACGAGCACGCCGTGTCCA 15851
Qy 2030 ProPheSerIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSer 2049
Db 15850 CCG-----------------------TGATCGCTGGGCCCTGCAGCCCGAGCACGTACG 15815
Qy 2050 ProGluGlyValGluProValSerProValSer---SerProSerLeuThrHisAspLys 2068
Db 15814 CCACACGGCCTGAGAGCACGCTCGCCGCGCTCCCGGTGCCCCGGT--------------- 15770
Qy 2069 GlyLeuProLysHisLeuGluGluLeuAspLysSerHisLeuGluGlyGluLeuArgPro 2088
Db 15769 AGCCGTCCAACGCCGCCAGGTCGCTGCCGTAAAGAGCACCGTAGTCCGAGC------CCA 15716
Qy 2089 LysGlnProGlyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgPro 2108
Db 15715 TCGAGCCCAGGTACACCCCGGTGGAGCTCTCGCTCAGCGCCGACGGGCGCACGCCAGCCC 15656
Qy 2109 LeuProGluSerGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGly 2128
Db 15655 GCTCCA---GTGCCTCCCACGCCGTCTCCA-----------GCACCAGCCGCTGCTGGG 15611
Qy 2129 HisGlnArgValValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThr 2148
Db 15610 GATCCATCGCCTGCGCCTCTCGGGGCG-------------------------ATATCC 15578
Qy 2149 ArgHisHisProGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGly 2168
Db 15577 CGAAGAAACCCGCATCGA---ACAGGTCGATGTCCCGCAGGAATCCACCCTCGCGCACGT 15521
Qy 2169 AlaSerCysProValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProPro 2188
Db 15520 AGGTCTTGCCCA-----------------------------------CCGCCTCCG 15500
Qy 2189 AspHisGlyAlaProAlaArgGlySerProHisSerGluGlyGlyLysArg--------- 2205
Db 15499 GATCGGGGTCGTAAATCGAAGGTGCATCCCAGCGCGCAGGGAAGGCCTCGATCGCATCTC 15440
Qy 2206 SerProGluProAsnLysThrSer---------------------------------- 2213
Db 15439 TCCCCTCCTCCAGGAGACGCCAGTAGTCTTCCGGCGTAGCGACGCCCCCCGGCAGCCGGC 15380
Qy 2214 ------------------------ValLeuGlyGlyGlyGluAspGlyIleGluProVal 2225
Db 15379 ACGCCATCGATACGATGGCGATGGGCCCTGAACGGAGATCCGAGTGGCTGGCCTGCAGCC 15320
Qy 2226 SerProProGluGlyMet------ThrGluProGlyHisSer------------------ 2237
Db 15319 TCTGCAACTCTTTCAATGCACGGGTCAACGCCTCACGAAGCGTAGGAGAGGCGGCGCTGC 15260
Qy 2238 ---------ArgSerAlaValTyrProLeuLeuTyrArg-----------------Asp 2248
Db 15259 TATTCATGGAGATCCTCTGAGTTCGAAATCATTTGCAAGGCTTGCAAGGCCAGCTGCTCG 15200
Qy 2249 GlyGluGlnThrGluProSerArgMetGlySer--------------------------- 2259
Db 15199 TCGGCGAGATCTTCGAAATCGGAAATGGGAACCGGAACCTCGACGGATGTTTCGGACGGC 15140
Qy 2260 ---LysSerProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSer 2278
Db 15139 CCCCGCAGCTCCAAGAGCTTGGGGAGGAGGCCTGAACGACGCACGAGACTGACGGAGAGG 15080
Qy 2279 AsnSerAlaMetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsn 2298
Db 15079 CTGAGCATGAACGCGCGAATTTCCTCGTCGGACATGGAGTCCGCTGTACGCAGCAGATTC 15020
Qy 2299 ArgAsnGluProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAla 2318
Db 15019 CGGCGCTCACCGTTTCCGAACTTTGACAGGAGGAAC---------TTCGCGATGTGCGTG 14969
Qy 2319 IleThrGly-----------------ThrGlyLeuMetThrTyrArgSerGlnAlaVal 2332

```
Db      14968 GCAGATGGATGGTCGAAAAGCAAAGTGGCAGGCAAGTTGCTTCCGACCAGCAAGGCGAGC 14909

Qy       2333 GlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaLeuMetGly 2352

Db      14908 CGATTGCGGACTTCGACGGCCATG---------------------------------- 14885

Qy       2353 LysTyrAspGlnTrpGluGluSerProProLeuSerAlaAsnAlaPheAsnProLeuAsn 2372

Db      14884 --------------AGCGAATCCATTCCGAGCTCCCTCAAGGGCTGGGCCGTAGGAATC 14840

Qy       2373 AlaSerAlaSerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHis 2392

Db      14839 GCGTCA---------CCGCGCTGAAGCCCGAGCACCGCGGCGACCTCGGCCCGCACAAGG 14789

Qy       2393 ThrLeuThrSerProGlyGlyGlyGlyLysAlaLysValSerGlyArgProSerSerArg 2412

Db      14788 TCGATGAGCGAGCTC-----------AGGCGCTCCGCCTCCGGCAGCGCCGAGAGGCGC 14741

Qy       2413 LysAlaLysSerProAlaProGlyLeuAlaSerGlyAsp------ArgProProSerVal 2430

Db      14740 TCGCGGAGCGTCGACGCCTCCTTCCTCGCGGCAGAGGACGCCTTGCGCAGCCCGGGGCGC 14681

Qy       2431 SerSerValHisSerGluGly--------------------------------AspCys 2439

Db      14680 ACCAGCGCGCGAAGCAGCGCCGGCAGCTCGCCGCTGGACTCCAGCCCACGCTGCAGCTGC 14621

Qy       2440 AsnArgArgThrProLeuThrAsnArgValTrpGluAspArgProSerSerAlaGlySer 2459

Db      14620 GCGAGATCCAGGTGCACTGGGACCAGGCTCGCTTCAGAGCGTGAGAGCGCGGCGTCCAGC 14561

Qy       2460 ThrProPheProTyrAsn-ProLeu------------------------------IleMe 2469

Db      14560 AGAGCGAGGCCCTCGTCGACCGATATCGGCACGAGCCCTGCGCGCCTGATGCGCGACAGT 14501

Qy       2469 tArgLeuGlnAlaGlyValMetAlaSerProProPro------ProGlyLeuProAla-- 2486

Db      14500 TCGGCCTCGCCCAGGTGCGCTGTCATCCCCACGCCCGCTTGCGCCCACAACCCCCACGCC 14441

Qy       2487 -GlySerGlyProLeuAla 2492

Db      14440 AGGCTCGTCGCCGCGAGCC 14422
```

```
RESULT 3
US-08-372-652-9
; Sequence 9, Application US/08372652
; Patent No. 5932699
;  GENERAL INFORMATION:
;    APPLICANT:  Moore, David
;    APPLICANT:  Seol, Wongi
;    APPLICANT:  Choi, Hueng-Sik
;    TITLE OF INVENTION:  RETINOID X RECEPTOR-INTERACTING
;    TITLE OF INVENTION:  POLYPEPTIDES AND RELATED MOLECULES AND METHODS
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street, Suite 3100
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/372,652
;      FILING DATE:  13-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/246001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1922 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-372-652-9

Alignment Scores:
Pred. No.:                9.5e-22        Length:          1922
Score:                    718.00         Matches:         212
Percent Similarity:       48.81%         Conservative:    76
Best Local Similarity:    35.93%         Mismatches:      214
Query Match:              5.43%          Indels:          88
DB:                       2              Gaps:            25

US-09-522-753-5 (1-2517) x US-08-372-652-9 (1-1922)

Qy       1975 ProSerLysGlySerGluProArgProLeuVal--------------------------- 1985

Db        209 CCAAGTGGCAAGGCCCAGCCTCATGCCTCAGTAGTGTATTCTGAGGCTGGGAAAGATAAA 268

Qy       1986 ------ProProValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsn------ 2001

Db        269 GGGCCTCCTCCAAAATCCAGATATGAGGAAGAGCTAAGGACCCGAGGGAAGACTACCATT 328

Qy       2002 LeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAspPro 2021

Db        329 ACTGCAGCTAACTTCATAGACGTGACCATCACCCGGCAAATTGCCTCGGACAAGGATGCG 388

Qy       2022 HisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSerLeu 2041

Db        389 AGGGAACGTGGCTCTCAAAGTTCAGACTCTTCT-----------------AGTAGCTTG 430

Qy       2042 GlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSerSer 2061

Db        431 TCTTCTCACAGGTATGAAACGGCTAGTGATGCCATTGAGGTGATAAGTCCCGCCAGCTCA 490

Qy       2062 ProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHis 2081

Db        491 CCTGCACCACCCCAGGAAAAGCCACAGGCCTATCAGCCAGACATGGTTAAGGCAAATCAA 550

Qy       2082 LeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAlaAla 2101

Db        551 GCAGAAAATGAGTCCACTCGACAGTATGAAGGTCCA----------------------- 586

Qy       2102 HisLeuProHisLeuArg--------------ProLeuProGluSerGlnProSerSer 2116

Db        587 ---CTGCATCATTATCGGTCCCAGCAGGAATCACCATCTCCACAGCAACAGCCACCACTG 643

Qy       2117 SerProLeuLeuGlnThr------AlaProGlyValLysGlyHisGlnArgValValThr 2134

Db        644 CCCCCATCTTCCCAGTCAGAGGGAATGGGACAGGTGCCCAGGACCCATCGACTGATCACA 703

Qy       2135 LeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisProGlnGln 2154

Db        704 CTTGCTGACCACATCTGTCAAATTATCACACAAGATTTTGCTAGAAAT------CAAGTT 757

Qy       2155 LeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAla------SerCysPro 2172

Db        758 CCCTCGCAGCCTTCTACTTCTACATTCCAAACTTCACCATCTGCTTTGTCATCCACACCT 817

Qy       2173 ValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAsp--------- 2189

Db        818 GTA---------AGAACTAAAACCTCAAGCCGCTACAGCCCAGAATCACAGTCTCAGACT 868

Qy       2190 ------HisGlyAlaProAla---ArgGlySerProHis---------------SerGlu 2201
```

```
Db        869 GTCTTGCATCCCAGACCAGGTCCTAGAGTCTCTCCAGAAAATCTTGTGGATAAATCCCGG 928

Qy       2202 GlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGly 2221

Db        929 GGAAGCAGGCCTGGAAAATCTCCAGAGAGGAGTCATATC---------CCATCAGAGCCC 979

Qy       2222 IleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAlaVal 2241

Db        980 TATGAGCCCATCTCCCCACCCCAAGGCCCTGCTGTG------CATGAGAAGCAGGACAGC 1033

Qy       2242 TyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSer 2261

Db       1034 ATGTTGCTCTTGTCACAGAGGGGAGTGGACCCTGCTGAGCAAAGGAGTGATTCTCGATCA 1093

Qy       2262 ProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAla 2281

Db       1094 CCAGGAAGTATAAGCTACTTGCCTTCATTCTTCACCAAGCTT---GAAAGCACATCACCC 1150

Qy       2282 MetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGlu 2301

Db       1151 ATGGTTAAATCAAAGAAACAGGAAATTTTTCGTAAGTTGAACTCTTCTGGTGGAGGTGAC 1210

Qy       2302 ProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGly 2321

Db       1211 TCTGATATGGCAGCTGCTCAGCCAGGAACAGAGATCTTCAATCTGCCAGCAGTTACCACA 1270

Qy       2322 ThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGly 2341

Db       1271 TCAGGTGCAGTGAGCTCAAGAAGCCATTCTTTTGCTGATCCCGCCAGT---AACCTTGGT 1327

Qy       2342 LeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGlu------ 2359

Db       1328 CTAGAAGACATCATCAGAAAGGCTCTCATGGGAAGTTTTGATGATAAAGTTGAAGATCAT 1387

Qy       2360 ------------SerProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAla 2375

Db       1388 GGTGTTGTCATGTCCCATCCTGTG---------GGCATTATGCCTGGTAGTGCCAGCACC 1438

Qy       2376 SerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThr 2395

Db       1439 TCAGTG--------------GTGACGAGCAGCGAGGCACGGAGAGATGAAGGGGAGCCA 1483

Qy       2396 SerProGlyGlyGly---GlyLysAlaLysValSerGlyArgProSerSerArgLysAla 2414

Db       1484 TCACCTCATGCAGGAGTATGCAAACCAAAGCTGATCAACAAATCAAACAGCAGGAAGTCT 1543

Qy       2415 LysSerProAlaProGlyLeuAla------SerGlyAspArgProProSerValSerSer 2432

Db       1544 AAATCTCCTATTCCTGGGCAAAGCTATTTAGGAACTGAAAGGCCTTCTTCTGTCTCCTCT 1603

Qy       2433 ValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAsp 2452

Db       1604 GTGCATTCAGAAGGTGATTACCACAGGCAGACACCA------GGATGGGCATGGGAAGAT 1657

Qy       2453 ArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGln 2472

Db       1658 CGGCCCTCTTCAACAGGTTCTACTCAGTTCCCTTACAACCCTCTGACCATACGGATG--- 1714

Qy       2473 AlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu--- 2491

Db       1715 ---------CTCAGCAGTACACCACCTACACAGATCGCATGCGCCCCATCTGCCATCACC 1765

Qy       2492 ---AlaGlyProHis---------HisAlaTrpAspGluGluProLysProLeuLeuCys 2507

Db       1766 CAAGCAGCTCCACATCAACAGAACCGCATCTGGGAGAGGGAGCCTGCCCCGCTCCTCTCA 1825

Qy       2508 SerGlnTyrGluThrLeuSerAspSerGlu 2517

Db       1826 GCGCAGTATGAGACACTGTCTGATAGTGAC 1855

RESULT 4
PCT-US95-16311-9
; Sequence 9, Application PC/TUS9516311
;  GENERAL INFORMATION:
;    APPLICANT:  Moore, David
;    APPLICANT:  Seol, Wongi
;    APPLICANT:  Choi, Hueng-Sik
;    TITLE OF INVENTION:  RETINOID X RECEPTOR-INTERACTING
;    TITLE OF INVENTION:  POLYPEPTIDES AND RELATED MOLECULES AND METHODS
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street, Suite 3100
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/16311
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,652
;      FILING DATE:  13-JAN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/246001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1922 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
PCT-US95-16311-9

Alignment Scores:
Pred. No.:              9.5e-22         Length:         1922
Score:                  718.00          Matches:        212
Percent Similarity:     48.81%          Conservative:   76
Best Local Similarity:  35.93%          Mismatches:     214
Query Match:            5.43%           Indels:         88
DB:                     5               Gaps:           25

US-09-522-753-5 (1-2517) x PCT-US95-16311-9 (1-1922)

Qy       1975 ProSerLysGlySerGluProArgProLeuVal-------------------------- 1985

Db        209 CCAAGTGGCAAGGCCCAGCCTCATGCCTCAGTAGTGTATTCTGAGGCTGGGAAAGATAAA 268

Qy       1986 ------ProProValSerGlyHisAlaThrIleAlaArgThrProAlaLysAsn------ 2001

Db        269 GGGCCTCCTCCAAAATCCAGATATGAGGAAGAGCTAAGGACCCGAGGGAAGACTACCATT 328

Qy       2002 LeuAlaProHisHisAlaSerProAspProProAlaProProAlaSerAlaSerAspPro 2021

Db        329 ACTGCAGCTAACTTCATAGACGTGACCATCACCCGGCAAATTGCCTCGGACAAGGATGCG 388

Qy       2022 HisArgGluLysThrGlnSerLysProPheSerIleGlnGluLeuGluLeuArgSerLeu 2041

Db        389 AGGGAACGTGGCTCTCAAAGTTCAGACTCTTCT----------------AGTAGCTTG 430

Qy       2042 GlyTyrHisGlySerSerTyrSerProGluGlyValGluProValSerProValSerSer 2061

Db        431 TCTTCTCACAGGTATGAAACGGCTAGTGATGCCATTGAGGTGATAAGTCCCGCCAGCTCA 490

Qy       2062 ProSerLeuThrHisAspLysGlyLeuProLysHisLeuGluGluLeuAspLysSerHis 2081
```

```
Db        491 CCTGCACCACCCCAGGAAAAGCCACAGGCCTATCAGCCAGACATGGTTAAGGCAAATCAA 550

Qy       2082 LeuGluGlyGluLeuArgProLysGlnProGlyProValLysLeuGlyGlyGluAlaAla 2101

Db        551 GCAGAAAATGAGTCCACTCGACAGTATGAAGGTCCA---------------------------- 586

Qy       2102 HisLeuProHisLeuArg---------------ProLeuProGluSerGlnProSerSer 2116

Db        587 ---CTGCATCATTATCGGTCCCAGCAGGAATCACCATCTCCACAGCAACAGCCACCACTG 643

Qy       2117 SerProLeuLeuGlnThr------AlaProGlyValLysGlyHisGlnArgValValThr 2134

Db        644 CCCCCATCTTCCCAGTCAGAGGGAATGGGACAGGTGCCCAGGACCCATCGACTGATCACA 703

Qy       2135 LeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisProGlnGln 2154

Db        704 CTTGCTGACCACATCTGTCAAATTATCACACAAGATTTTGCTAGAAAT------CAAGTT 757

Qy       2155 LeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAla------SerCysPro 2172

Db        758 CCCTCGCAGCCTTCTACTTCTACATTCCAAACTTCACCATCTGCTTTGTCATCCACACCT 817

Qy       2173 ValLeuAspLeuArgArgProProSerAspLeuTyrLeuProProProAsp--------- 2189

Db        818 GTA---------AGAACTAAAACCTCAAGCCGCTACAGCCCAGAATCACAGTCTCAGACT 868

Qy       2190 ------HisGlyAlaProAla---ArgGlySerProHis--------------SerGlu 2201

Db        869 GTCTTGCATCCCAGACCAGGTCCTAGAGTCTCTCCAGAAAATCTTGTGGATAAATCCCGG 928

Qy       2202 GlyGlyLysArgSerProGluProAsnLysThrSerValLeuGlyGlyGlyGluAspGly 2221

Db        929 GGAAGCAGGCCTGGAAAATCTCCAGAGAGGAGTCATATC---------CCATCAGAGCCC 979

Qy       2222 IleGluProValSerProProGluGlyMetThrGluProGlyHisSerArgSerAlaVal 2241

Db        980 TATGAGCCCATCTCCCCACCCCAAGGCCCTGCTGTG------CATGAGAAGCAGGACAGC 1033

Qy       2242 TyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSerArgMetGlySerLysSer 2261

Db       1034 ATGTTGCTCTTGTCACAGAGGGGAGTGGACCCTGCTGAGCAAAGGAGTGATTCTCGATCA 1093

Qy       2262 ProGlyAsnThrSerGlnProProAlaPhePheSerLysLeuThrGluSerAsnSerAla 2281

Db       1094 CCAGGAAGTATAAGCTACTTGCCTTCATTCTTCACCAAGCTT---GAAAGCACATCACCC 1150

Qy       2282 MetValLysSerLysLysGlnGluIleAsnLysLysLeuAsnThrHisAsnArgAsnGlu 2301

Db       1151 ATGGTTAAATCAAAGAAACAGGAAATTTTTCGTAAGTTGAACTCTTCTGGTGGAGGTGAC 1210

Qy       2302 ProGluTyrAsnIleSerGlnProGlyThrGluIlePheAsnMetProAlaIleThrGly 2321

Db       1211 TCTGATATGGCAGCTGCTCAGCCAGGAACAGAGATCTTCAATCTGCCAGCAGTTACCACA 1270

Qy       2322 ThrGlyLeuMetThrTyrArgSerGlnAlaValGlnGluHisAlaSerThrAsnMetGly 2341

Db       1271 TCAGGTGCAGTGAGCTCAAGAAGCCATTCTTTTGCTGATCCCGCCAGT---AACCTTGGT 1327

Qy       2342 LeuGluAlaIleIleArgLysAlaLeuMetGlyLysTyrAspGlnTrpGluGlu------ 2359

Db       1328 CTAGAAGACATCATCAGAAAGGCTCTCATGGGAAGTTTTGATGATAAAGTTGAAGATCAT 1387

Qy       2360 -----------SerProProLeuSerAlaAsnAlaPheAsnProLeuAsnAlaSerAla 2375

Db       1388 GGTGTTGTCATGTCCCATCCTGTG---------GGCATTATGCCTGGTAGTGCCAGCACC 1438

Qy       2376 SerLeuProAlaAlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeuThr 2395

Db       1439 TCAGTG--------------GTGACGAGCAGCGAGGCACGGAGAGATGAAGGGGAGCCA 1483

Qy       2396 SerProGlyGlyGly---GlyLysAlaLysValSerGlyArgProSerSerArgLysAla 2414

Db       1484 TCACCTCATGCAGGAGTATGCAAACCAAAGCTGATCAACAAATCAAACAGCAGGAAGTCT 1543

Qy       2415 LysSerProAlaProGlyLeuAla------SerGlyAspArgProProSerValSerSer 2432

Db       1544 AAATCTCCTATTCCTGGGCAAAGCTATTTAGGAACTGAAAGGCCTTCTTCTGTCTCCTCT 1603

Qy       2433 ValHisSerGluGlyAspCysAsnArgArgThrProLeuThrAsnArgValTrpGluAsp 2452

Db       1604 GTGCATTCAGAAGGTGATTACCACAGGCAGACACCA------GGATGGGCATGGGAAGAT 1657

Qy       2453 ArgProSerSerAlaGlySerThrProPheProTyrAsnProLeuIleMetArgLeuGln 2472

Db       1658 CGGCCCTCTTCAACAGGTTCTACTCAGTTCCCTTACAACCCTCTGACCATACGGATG--- 1714

Qy       2473 AlaGlyValMetAlaSerProProProProGlyLeuProAlaGlySerGlyProLeu--- 2491

Db       1715 ---------CTCAGCAGTACACCACCTACACAGATCGCATGCGCCCCATCTGCCATCACC 1765

Qy       2492 ---AlaGlyProHis---------HisAlaTrpAspGluGluProLysProLeuLeuCys 2507

Db       1766 CAAGCAGCTCCACATCAACAGAACCGCATCTGGGAGAGGGAGCCTGCCCCGCTCCTCTCA 1825

Qy       2508 SerGlnTyrGluThrLeuSerAspSerGlu 2517

Db       1826 GCGCAGTATGAGACACTGTCTGATAGTGAC 1855
```

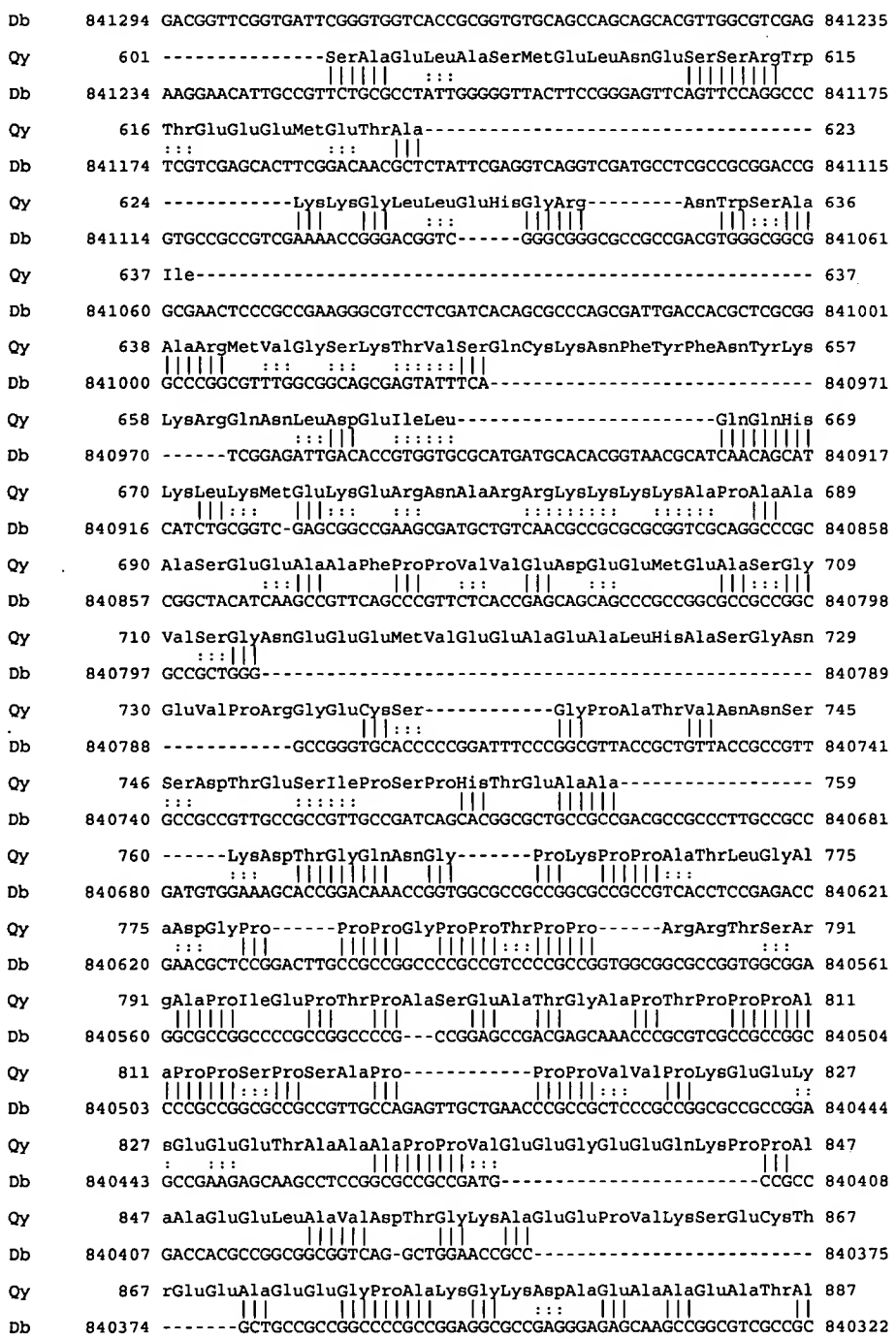
RESULT 5
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

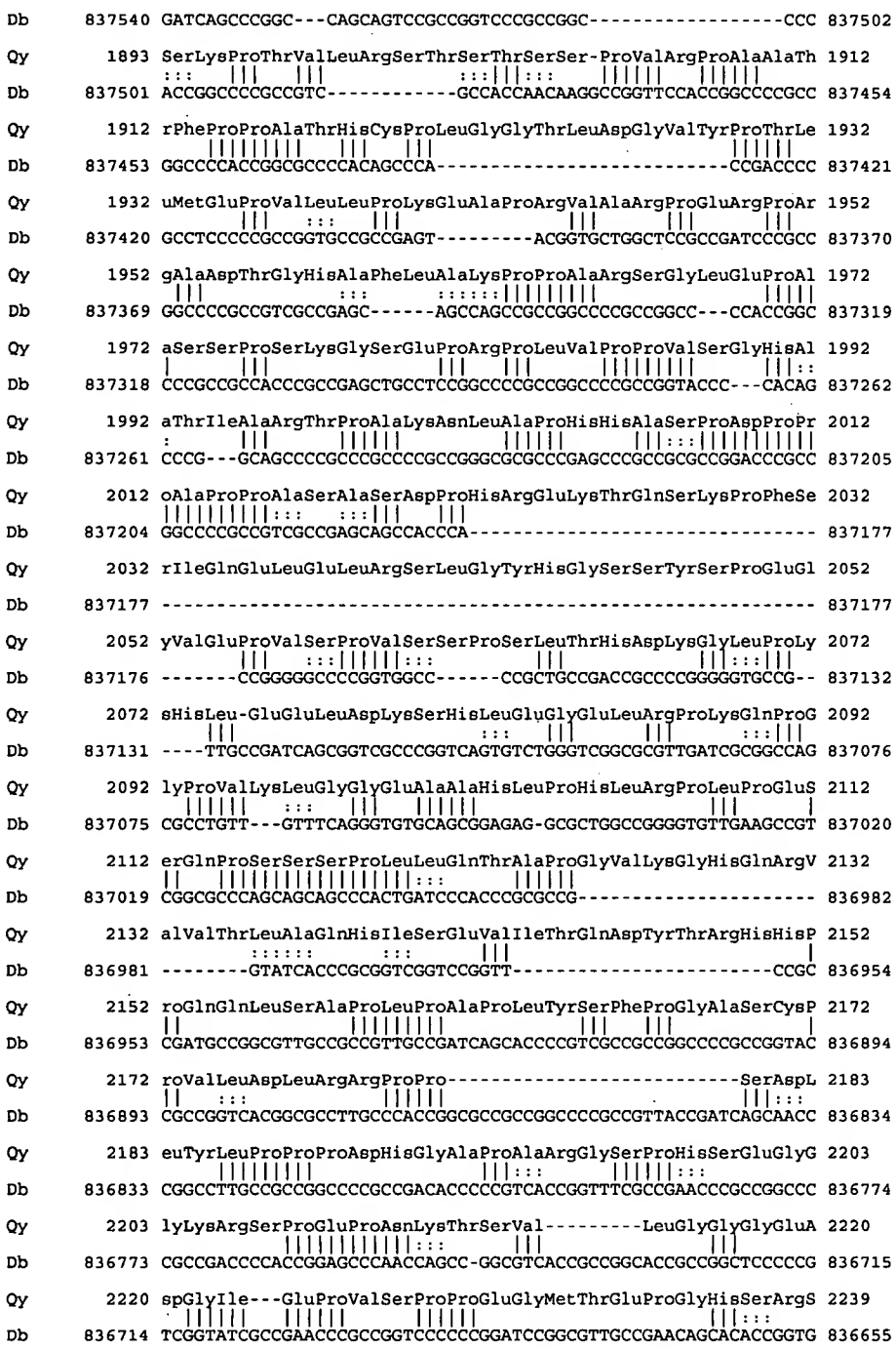
Alignment Scores:

Pred. No.:	3.28e-17	Length:	4411529
Score:	689.00	Matches:	526
Percent Similarity:	33.44%	Conservative:	207
Best Local Similarity:	24.00%	Mismatches:	872
Query Match:	5.21%	Indels:	599
DB:	3	Gaps:	91

US-09-522-753-5 (1-2517) x US-09-103-840A-1 (1-4411529)

```
Qy        546 LeuLysGluLysThrAspAspThrSerGlyGluAspAsnAspGluLysGluAlaValAla 565

Db     841474 GTTCGTCGGAAGGTGATGGGGCTGGGCGGTGACGGCTTCGACGAATGCGAATGCCTCTGC 841415

Qy        566 SerLysGlyArgLysThrAlaAsnSer----------------GlnGlyArgArg--- 578

Db     841414 TCGCGGTGACGGAATCTCGAAGATGCGTCGATTCGTTGCCAGCCGGAGGAACGACGCCCA 841355

Qy        579 -----------------LysGlyArgIleThrArgSerMetAlaAsnGluAlaAsnSer 592

Db     841354 TACCAGGTTCGGCACTGTGAAGGGGTCGTCAGCCGCGAGCAGTCGATCGAACCAGGGGCG 841295

Qy        593 Glu--------------------GluAlaIleThrProGlnGln--------------- 600
```


Db 841294 GACGGTTCGGTGATTCGGGTGGTCACCGCGGTGTGCAGCCAGCAGCACGTTGGCGTCGAG 841235
Qy 601 ----------SerAlaGluLeuAlaSerMetGluLeuAsnGluSerSerArgTrp 615
Db 841234 AAGGAACATTGCCGTTCTGCGCCTATTGGGGGTTACTTCCGGGAGTTCAGTTCCAGGCCC 841175
Qy 616 ThrGluGluGluMetGluThrAla---------- 623
Db 841174 TCGTCGAGCACTTCGGACAACGCTCTATTCGAGGTCAGGTCGATGCCTCGCCGCGGACCG 841115
Qy 624 ----------LysLysGlyLeuLeuGluHisGlyArg---------AsnTrpSerAla 636
Db 841114 GTGCCGCCGTCGAAAACCGGGACGGTC------GGGCGGGCGCCGCCGACGTGGGCGGCG 841061
Qy 637 Ile---------- 637
Db 841060 GCGAACTCCCGCCGAAGGGCGTCCTCGATCACAGCGCCCAGCGATTGACCACGCTCGCGG 841001
Qy 638 AlaArgMetValGlySerLysThrValSerGlnCysLysAsnPheTyrPheAsnTyrLys 657
Db 841000 GCCCGGCGTTTGGCGGCAGCGAGTATTTCA---------- 840971
Qy 658 LysArgGlnAsnLeuAspGluIleLeu----------GlnGlnHis 669
Db 840970 ------TCGGAGATTGACACCGTGGTGCGCATGATGCACACGGTAACGCATCAACAGCAT 840917
Qy 670 LysLeuLysMetGluLysGluArgAsnAlaArgArgLysLysLysLysAlaProAlaAla 689
Db 840916 CATCTGCGGTC-GAGCGGCCGAAGCGATGCTGTCAACGCCGCGCGCGGTCGCAGGCCCGC 840858
Qy 690 AlaSerGluGluAlaAlaPheProProValValGluAspGluGluMetGluAlaSerGly 709
Db 840857 CGGCTACATCAAGCCGTTCAGCCCGTTCTCACCGAGCAGCAGCCCGCCGGCGCCGCCGGC 840798
Qy 710 ValSerGlyAsnGluGluGluMetValGluGluAlaGluAlaLeuHisAlaSerGlyAsn 729
Db 840797 GCCGCTGGG---------- 840789
Qy 730 GluValProArgGlyGluCysSer----------GlyProAlaThrValAsnAsnSer 745
Db 840788 ----------GCCGGGTGCACCCCCGGATTTCCCGGCGTTACCGCTGTTACCGCCGTT 840741
Qy 746 SerAspThrGluSerIleProSerProHisThrGluAlaAla---------- 759
Db 840740 GCCGCCGTTGCCGCCGTTGCCGATCAGCACGGCGCTGCCGCCGACGCCGCCCTTGCCGCC 840681
Qy 760 ------LysAspThrGlyGlnAsnGly-------ProLysProProAlaThrLeuGlyAl 775
Db 840680 GATGTGGAAAGCACCGGACAAACCGGTGGCGCCGCCGGCGCCGCCGTCACCTCCGAGACC 840621
Qy 775 aAspGlyPro------ProProGlyProProThrProPro------ArgArgThrSerAr 791
Db 840620 GAACGCTCCGGACTTGCCGCCGGCCCCGCCGTCCCCGCCGGTGGCGGCGCCGGTGGCGGA 840561
Qy 791 gAlaProIleGluProThrProAlaSerGluAlaThrGlyAlaProThrProProProAl 811
Db 840560 GGCGCCGGCCCCGCCGGCCCCG---CCGGAGCCGACGAGCAAACCCGCGTCGCCGCCGGC 840504
Qy 811 aProProSerProSerAlaPro----------ProProValValProLysGluGluLy 827
Db 840503 CCCGCCGGCGCCGCCGTTGCCAGAGTTGCTGAACCCGCCGCTCCCGCCGGCGCCGCCGGA 840444
Qy 827 sGluGluGluThrAlaAlaAlaProProValGluGluGlyGluGluGlnLysProProAl 847
Db 840443 GCCGAAGAGCAAGCCTCCGGCGCCGCCGATG----------CCGCC 840408
Qy 847 aAlaGluGluLeuAlaValAspThrGlyLysAlaGluGluProValLysSerGluCysTh 867
Db 840407 GACCACGCCGGCGGCGGTCAG-GCTGGAACCGCC---------- 840375
Qy 867 rGluGluAlaGluGluGlyProAlaLysGlyLysAspAlaGluAlaAlaGluAlaThrAl 887
Db 840374 -------GCTGCCGCCGGCCCCGCCGGAGGCGCCGAGGGAGAGCAAGCCGGCGTCGCCGC 840322
Qy 887 aGluGlyAlaLeuLysAlaGluLysLysGluGlyGlySerGlyArgAlaThrThrAlaLy 907
Db 840321 CGGCCCCGCCGGCCCCTCCGGAAACTCCGGCGGCGGTGCTATGGCTGCCGGCGCCGCCGG 840262
Qy 907 sSerSerGlyAlaProGlnAspSerAspSerSerAlaThrCysSerAlaAspGluValAs 927
Db 840261 TGCCGCCGGCGCCGAACAGCCCGCCGTTGCCGCCCGCCCCGCCGGCGCCGCCGGCGAGGG 840202
Qy 927 pGluAlaGluGlyGlyAspLysAsnArgLeuLeuSerProArgProSerLeuLeuThrPr 947
Db 840201 TGCTTCCCCCGCCGGCCCCGCCGGTCCCGCCGGCGCCGAACAGCCCGCCGTTCCCACCTG 840142
Qy 947 oThrGlyAspProArgAlaAsnAlaSerProGlnLysProLeuAspLeuLysGlnLeuLy 967
Db 840141 CCCCGCCGATCCCGCCAGCCCCAGCGGGGCTGCCAAACCCACCGG-GCCCGCCGGCTC-- 840085
Qy 967 sGlnArgAlaAlaAlaIleProProIleGlnValThrLysValHisGluProProArgGl 987
Db 840084 -CCCCGTTG-GCGAACAGCCCGCCGTTA----------CCGCCGGCCCC 840048
Qy 987 uAspAlaAlaProThrLysProAlaProProAlaProProProProGlnAsnLeuGlnPr 1007
Db 840047 GCCATTGCCACCGACCGCGCCGGCACCACCGGCCCCGCCG---------- 840008
Qy 1007 oGluSerAspAlaProGlnGlnProGlySerSerProArgGlyLysSerArgSerProAl 1027
Db 840007 ----------GCCCCGGCGGCGCCCAAGAGCGCACCG---------GCGTTGCCGCCGGC 839967
Qy 1027 aProProAlaAspLysGluAlaPheAlaAlaGluAlaGlnLysLeuProGlyAspProPr 1047
Db 839966 CCCGCCATTGCCACCGACCGCGCCGGCACCACCGCTTCCGCCGGCCCCGCCGGAGCCGAA 839907
Qy 1047 oCysTrpThrSerGlyLeuProPheProValProProArgGluValIleLysAlaSerPr 1067
Db 839906 GAGCAAGGCGGCTCCGCCGGCCCCGCCAATCCCGCCGAGGCCGGTGGCGCCGGGCCCGCC 839847
Qy 1067 oHisAlaProAspProSerAlaPheSerTyrAlaProProGlyHisProLeuProLeuGl 1087
Db 839846 GCTGCCCCCGGTGCCGCCGGTGCCGAACAAGATCCCGCCGGCCCCGCCGGCCCCGCCGGC 839787
Qy 1087 y------LeuHisAspThrAlaArgProValLeuProArgProProThrIleSerAsnPr 1105
Db 839786 CCCGCCGTTGACGCCGCCGATCGTTCCCTGCCCGCCGTTTCCGCCGTTGCCGAACAACCA 839727
Qy 1105 oProProLeuIleSerSerAlaLysHisProSerValLeuGluArgGlnIleGlyAlaIl 1125
Db 839726 GCCGCCGGCTCCGCCGTCAGCCCCGGTCCCCGGAGTGCC----------GTTGGCGCCGT 839677
Qy 1125 eSerGln----------GlyMetSerValGlnLeuHisValProTyrSerGluHisAl 1141
Db 839676 TGCCGATCAGCGGACGCCCGGTGAGCGTCTGGAAGGGGTCGTTCACCATATTGATCACGT 839617
Qy 1141 a---------- 1141
Db 839616 CCTGCTGCAGGGTGTGCAGGGGCGAGGTGCTGGCGGGGGCCGTAAAGCCGTCCAGGCCCA 839557
Qy 1142 -LysAlaProValGlyProValThrMetGlyLeuProLeuProMetAspProLysLysLe 1161
Db 839556 ACAGCACCCCCCCGGTACCGCCGATGCCGGCCTTGCCCAGGGTCGCGCCGGTCCCGCCGC 839497
Qy 1161 uAlaProPheSerGlyValLysGlnGluGlnLeuSerProArgGlyGlnAlaGlyProPr 1181
Db 839496 TGCCGCCGTTGCCGCCGTTGCCGATCAACACGCCGGTGCCGCCGGCCCCGCCGGGCCCAC 839437
Qy 1181 oGluSerLeuGlyValProThrAla---GlnGluAlaSerValLeuArgGlyThrAlaLe 1200
Db 839436 C----------GTCCCCGCCGCCGGTGCTCGCGCCGCCGTTGCCGCCGTTGCCGCCGT 839389
Qy 1200 uGlySerValProGlyGlySerIleThrLys---------GlyIleProSerThrArgVa 1217
Db 839388 TGCCGATCAGCCCGGGGGCCCCGCCGGCCCCGCCGGCGGCGCTGTCCCCGACGCTTTTAC 839329


Db 837540 GATCAGCCCGGC---CAGCAGTCCGCCGGTCCCGCCGGC----------CCC 837502
Qy 1893 SerLysProThrValLeuArgSerThrSerThrSerSer-ProValArgProAlaAlaTh 1912
Db 837501 ACCGGCCCCGCCGTC----------GCCACCAACAAGGCCGGTTCCACCGGCCCCGCC 837454
Qy 1912 rPheProProAlaThrHisCysProLeuGlyGlyThrLeuAspGlyValTyrProThrLe 1932
Db 837453 GGCCCCACCGGCGCCCCACAGCCCA----------CCGACCCC 837421
Qy 1932 uMetGluProValLeuLeuProLysGluAlaProArgValAlaArgProGluArgProAr 1952
Db 837420 GCCTCCCCCGCCGGTGCCGCCGAGT---------ACGGTGCTGGCTCCGCCGATCCCGCC 837370
Qy 1952 gAlaAspThrGlyHisAlaPheLeuAlaLysProProAlaArgSerGlyLeuGluProAl 1972
Db 837369 GGCCCCGCCGTCGCCGAGC------AGCCAGCCGCCGGCCCCGCCGGCC---CCACCGGC 837319
Qy 1972 aSerSerProSerLysGlySerGluProArgProLeuValProProValSerGlyHisAl 1992
Db 837318 CCCGCCGCCACCCGCCGAGCTGCCTCCGGCCCCGCCGGCCCCGCCGGTACCC---CACAG 837262
Qy 1992 aThrIleAlaArgThrProAlaLysAsnLeuAlaProHisHisAlaSerProAspProPr 2012
Db 837261 CCCG---GCAGCCCCGCCCGCCCCGCCGGGCGCGCCCGAGCCCGCCGCGCCGGACCCGCC 837205
Qy 2012 oAlaProProAlaSerAlaSerAspProHisArgGluLysThrGlnSerLysProPheSe 2032
Db 837204 GGCCCCGCCGTCGCCGAGCAGCCACCCA---------- 837177
Qy 2032 rIleGlnGluLeuGluLeuArgSerLeuGlyTyrHisGlySerSerTyrSerProGluGl 2052
Db 837177 ---------- 837177
Qy 2052 yValGluProValSerProValSerSerProSerLeuThrHisAspLysGlyLeuProLy 2072
Db 837176 -------CCGGGGGCCCCGGTGGCC------CCGCTGCCGACCGCCCCGGGGGTGCCG-- 837132
Qy 2072 sHisLeu-GluGluLeuAspLysSerHisLeuGluGlyGluLeuArgProLysGlnProG 2092
Db 837131 ----TTGCCGATCAGCGGTCGCCCGGTCAGTGTCTGGGTCGGCGCGTTGATCGCGGCCAG 837076
Qy 2092 lyProValLysLeuGlyGlyGluAlaAlaHisLeuProHisLeuArgProLeuProGluS 2112
Db 837075 CGCCTGTT---GTTTCAGGGTGTGCAGCGGAGAG-GCGCTGGCCGGGGTGTTGAAGCCGT 837020
Qy 2112 erGlnProSerSerSerProLeuLeuGlnThrAlaProGlyValLysGlyHisGlnArgV 2132
Db 837019 CGGCGCCCAGCAGCAGCCCACTGATCCCACCCGCGCCG---------- 836982
Qy 2132 alValThrLeuAlaGlnHisIleSerGluValIleThrGlnAspTyrThrArgHisHisP 2152
Db 836981 --------GTATCACCCGCGGTCGGTCCGGTT----------CCGC 836954
Qy 2152 roGlnGlnLeuSerAlaProLeuProAlaProLeuTyrSerPheProGlyAlaSerCysP 2172
Db 836953 CGATGCCGGCGTTGCCGCCGTTGCCGATCAGCACCCCGTCGCCGCCGGCCCCGCCGGTAC 836894
Qy 2172 roValLeuAspLeuArgArgProPro----------SerAspL 2183
Db 836893 CGCCGGTCACGGCGCCTTGCCCACCGGCGCCGCCGGCCCCGCCGTTACCGATCAGCAACC 836834
Qy 2183 euTyrLeuProProProAspHisGlyAlaProAlaArgGlySerProHisSerGluGlyG 2203
Db 836833 CGGCCTTGCCGCCGGCCCCGCCGACACCCCCGTCACCGGTTTCGCCGAACCCGCCGGCCC 836774
Qy 2203 lyLysArgSerProGluProAsnLysThrSerVal---------LeuGlyGlyGlyGluA 2220
Db 836773 CGCCGACCCCACCGGAGCCCAACCAGCC-GGCGTCACCGCCGGCACCGCCGGCTCCCCCG 836715
Qy 2220 spGlyIle---GluProValSerProProGluGlyMetThrGluProGlyHisSerArgS 2239
Db 836714 TCGGTATCGCCGAACCCGCCGGTCCCCCCGGATCCGGCGTTGCCGAACAGCACACCGGTG 836655
Qy 2239 erAlaValTyrProLeuLeuTyrArgAspGlyGluGlnThrGluProSer---------- 2255
Db 836654 TTGCCGCCGGCCCC----------ACCGGCCCCACCGTCACCACCT 836619
Qy 2256 --------ArgMetGlySerLysSerProGlyAsnThrSerGlnProProAlaPhePheS 2273
Db 836618 AGGCCGGCTCCGCCGGATCCACCGGTGCCGCCGGCACCGAACAGCAGCCCGGCGTTGCCG 836559
Qy 2273 erLysLeuThrGluSerAsnSerAlaMetValLysSerLysLysGlnGluIleAsnLysL 2293
Db 836558 CCAGCCCCACCGACCCCACC----------GTTGTTGAGA 836529
Qy 2293 ysLeuAsnThrHisAsnArgAsnGluProGluTyrAsnIleSerGlnProGly------- 2310
Db 836528 AAGCCACGTCCGCCGGTCCCGCCGGCCCCGCCCGGCCCGGCCAGCAGCCCGGCCAGCAGC 836469
Qy 2311 --ThrGluIlePheAsnMetProAlaIleThrGlyThrGlyLeuMetThrTyrArgSerG 2330
Db 836468 CCACCGGCTCCGCCAGTTCCCCCGACCCCACCGGTGACGGCGC-CTAGCCCGCCGGCGCC 836410
Qy 2330 lnAlaValGlnGluHisAlaSerThrAsnMetGlyLeuGluAlaIleIleArgLysAlaL 2350
Db 836409 GCCGTGCCCGCCAGCAC---------- 836393
Qy 2350 euMetGlyLysTyrAspGlnTrpGluGluSerProProLeuSer-AlaAsnAlaPheAsn 2369
Db 836392 ----------CCAACAGCCCGCCGGCCCCACCAAC 836368
Qy 2370 ProLeuAsnAlaSerAlaSerLeuPro----------Ala 2379
Db 836367 CCCGCCGGCCCCACCGGCGCCGGCCCCGAGCCCGCCGACCCCGCCGACCCCACCGGTGCC 836308
Qy 2380 AlaMetProIleThrAlaAlaAspGlyArgSerAspHisThrLeu---------- 2394
Db 836307 CAACAACCAGCCACCTGATCCACCGGCACCCCCGGCCTCGCCATCACCTACCGTGGAACT 836248
Qy 2395 ----------ThrSerProGlyGly----------GlyGlyLysAlaLysValSer 2406
Db 836247 CCCGCCGGCCCCACCAGCCCCGCCGGTGCCGAACAACCCGGCGGCCCCGCCGGCCCCGCC 836188
Qy 2407 -GlyArgProSerSerArgLysAlaLysSerProAlaProGlyLeuAlaSerGlyAspAr 2426
Db 836187 GGGCAGGCCCGAGCCAGCGGCGGCGGACCCGCCGGCCCCACCATTACCGAGC---AACCA 836131
Qy 2426 gProProSerValSerSerValHisSerGluGlyAspCysAsnArgArgThrProLeuTh 2446
Db 836130 CCCGCCCGGCGCACCATTGGCCCCCGTGCCCGGG----------GCCCCGTTGGC 836086
Qy 2446 rAsnArgValTrpGluAspArgPro

This Page Blank (uspto)